IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: Stefek, *et al* **Docket No.:** BARR0005

Serial No. : 09/895,605 **Art Unit:** 2121

Filed: 29 June 2001 **Examiner:** Booker, Kevin E.

Title: AN INTEGRATIVE METHOD FOR MODELING MULTIPLE ASSET CLASSES

Assistant Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Declaration of prior invention to overcome cited patent or publication pursuant to 37 CFR 1.131

1. My name is Daniel Stefek.

2. I have reviewed the following document cited by the Examiner: Labe, JR. *et al* (U.S. Publication Number 2002/0091605).

3. The claimed subject matter of my invention is not what is claimed in the above cited reference.

4. The earliest filing date of the above mentioned reference is November 1, 2000. The conception of the claimed subject matter of my invention occurred prior to the specified date of said reference and was coupled with due diligence from prior to said reference date to the filing of the above mentioned application. In support of this, I have attached the document, Exhibit A below.

Exhibit A
"An Integrative Approach to Modeling the World Equity Market" dated June 23, 2000. Indicates that the concept of the invention was at least on or before June 23, 2000.

5. I am one of the inventors that used the attached document shown in Exhibit A. Exhibit A contains a description of the invention and was created on or before June 23, 2000, showing a conception date of the invention prior to the effective date of said reference.

6. The above-cited application was subsequently filed on June 29, 2001.

7. The above supporting facts show the conception of my invention prior to the effective date of said reference coupled with due diligence from prior to said date to the filing of the above-cited application.

8. I herein acknowledge that willful false statements and the like are punishable by fine or imprisonment, or both (18 U.S.C. 1001) and may jeopardize the validity of my application or any patent issuing thereon. All statements made of my own knowledge are true and all statements made on information or belief are believed to be true.

Daniel Stefek

Daniel Stefek,
Declarant

3/15/04

Date

#6

<u>IN THE UNITED STATES PATENT AND TRADEMARK OFFICE</u>

In re Application of: Stefek, *et al* **Docket No.:** BARR0005

Serial No. : 09/895,605 **Art Unit:** 2121

Filed: 29 June 2001 **Examiner:** Booker, Kevin E.

Title: AN INTEGRATIVE METHOD FOR MODELING MULTIPLE ASSET CLASSES

March 15, 2004

Assistant Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

RECEIVED
MAR 2 2 2004
Technology Center 2100

Declaration of prior invention to overcome cited patent or publication pursuant to 37 CFR 1.131

1. My name is Noel Johnson, Esq.

2. I am the Associate General Counsel for Barra, Inc. and oversee the development and management of Barra's patent portfolio. I have reviewed Exhibits B and C and hereby acknowledge and verify that the documents are correspondences directly related to the due process and filing of the present patent application, as follows:

<u>Exhibit B</u>
Letter to the internal legal counsel department at Barra from outside patent counsel dated January 29, 2001 indicating a draft of the patent application was enclosed for the inventors' review.
Such document indicates that a draft of the patent application was prepared on or before January 29, 2001.

Exhibit C

Letter to the internal legal counsel department at Barra from outside patent counsel dated April 24, 2001 indicating a draft of the patent application was enclosed for the inventors' review.

Such document indicates that another draft of the patent application was prepared on or before April 24, 2001.

3. I herein acknowledge that willful false statements and the like are punishable by fine or imprisonment, or both (18 U.S.C. 1001) and may jeopardize the validity of my application or any patent issuing thereon. All statements made of my own knowledge are true and all statements made on information or belief are believed to be true.

Noel Johnson, Esq.,
Declarant

3/15/04
Date

An Integrative Approach to Modeling the World Equity Market

Pei Chen, Fati Hemmati, Nicolo G. Torre, Ph.D.

June 23, 2000

Contents

"All my fortunes are at sea"
The Merchant of Venice
William Shakespeare

1 Introduction

Our modest goal in the following is to develop a new approach to modeling global equity risk. We shall begin by describing prior art in this domain. We shall go on to an account of our preliminary research, leading up to a choice of modeling strategy. Then we shall describe the model estimation in detail. Finally we close with a comparison of risk analysis conducted in the new model and in GEM.

2 Prior Art

2.1 Fundamental Concepts

A standard approach to characterizing financial risk is to measure the variance of the return series. Let $r_i(t)$ be the return to asset i in period t and define the asset-to-asset covariance matrix Ω by

$$\Omega_{ij}(t) = \mathrm{cov}(r_i(t), r_j(t))$$

where cov( ) is the covariance operator. Suppose h is the vector whose component h_i gives the fraction of wealth invested in asset i by a particular portfolio. Then the variance of the portfolio return is given by

$$h^t\Omega h$$

More generally if h_1 and h_2 are vectors defining distinct portfolios, then the covariance of the returns to those portfolios is given by $h_1^t\Omega h_2$. This remark reduces the risk analysis of a portfolio to the problem of determining a good estimate $\hat{\Omega}$ of the asset-by-asset covariance matrix Ω.

The simplest estimator $\hat{\Omega}$ is the historical covariance matrix

$$\Omega_{ij}^{hist}(t) = \mathrm{cov}(\{r_i(u)\}_{u=1}^t, \{r_j(v)\}_{v=1}^t)$$

The statistical properties of this estimator depend crucially on two parameters: the number of time periods T entering the estimate and the number of assets N covered by the estimate. Let T denote the total length of time over which

returns are observed and let Δ denote the observation frequency. Then $t = T/\Delta$. In general T is limited by two circumstances. First, assets have finite lives. Second, the economy itself is evolving and this evolution limits the relevance of data from the distant past. A practical choice for T is about five years. The choices of observation frequency Δ is also constrained. An empirical finding is that the return structure of the market differs depending on the time scale of observation. This point is illustrated in figure one, which shows how the return structure of different horizons is fit by models tuned to the daily and monthly time scales respectively. For the risk horizons of interest in a portfolio management context (from one quarter to a few years) a rule of thumb is that the observation frequency Δ is best set at a one month horizon. Taking T at five years and Δ at one month results in the number of periods t being 60. The choice of 60 is not a hard number, but it represents a reasonable and necessary compromise.

With this preamble we are now able to discuss the statistical properties of the historical estimator. Let Ω_{ij} represent the true value of the covariance and let $\hat{\Omega}_{ij}$ be its estimate. The true correlation ρ_{ij} between two assets is

$$\rho_{ij} = \Omega_{ij} / \sqrt{\Omega_{ij}\ \Omega_{ij}}$$

For normally distributed returns the quantity

$$\hat{\Omega}_{ij}\ /\ \Omega_{ij}\ \Omega_{ij}$$

is approximately normally distributed with mean ρ and standard deviation $\sqrt{(1+\rho^2)/t}$. Taking the number of assets $N = 1000$, plugging in typical values for the other terms and treating the cells of the covariance matrix as if they were independent draws from a normal distribution (which of course they are not) one finds that approximately 600 of the cells will be measured with the wrong sign. While this analysis is clearly limited, it shows that $\hat{\Omega}$ must be approached with caution. Let h be an arbitrary portfolio. Then the quantity

$$t\ \ h^t\hat{\Omega}h\ /\ h^t\Omega h$$

will be distributed as a chi square variable χ^2 with t degrees of freedom. The variance of chi_t^2 is $2t$, so for t=60 we find that the standard deviation of $h^t\hat{\Omega}h\ /\ h^t\Omega h$ is approximately 3%. On the other hand, this ratio will fall outside the range [0.7, 1.3] approximately 10% of the time. For portfolios constructed by mean-variance optimization expected "worse case" performance of the estimator is of greater relevance than average performance. A formal analysis of this situation requires a precise definition of worst case and is beyond our present scope. A rough indication of how badly the estimator may perform is given by considering the ratio of the eigenvalues?? of $\hat{\Omega}$ to the eigenvalues of Ω.

Let $\{\lambda_i\}$ be the true eigenvalues and $\{\hat{\lambda}_i\}$ be the estimated values. Then the quantities $\hat{\lambda}_i \ / \ \lambda_i$ are independently normally distributed with mean one and standard deviation $\sqrt{2/t}$. Among N distinct eigenvalues let

$$M_N = \max\{\hat{\lambda}_i \ / \ \lambda_i\}$$

Then we can give the expected value of M_N and an upper bound α such that $M_N > \alpha_N$ only 5% of the time as

N	E[M_N]	α_N
50	1.42	1.50
100	1.47	1.55
500	1.56	1.63
1000	1.60	1.65

A second difficulty with the historical covariance matrix is that if $N > T$ the matrix will fail to be positive definite. This technical term indicates that it is possible to find portfolios h which appear to be risk free, i.e. $h^t \Omega^{hist} \ h = 0$. In fact since $N > T$ we see that we can find non-zero linear combinations h_i such that

$$\sum_i h_i \ r_i(t) = 0 \tag{1}$$

For such h_i to define a portfolio it is necessary only that either

$$\sum_i h_i = 1 \tag{2}$$

or

$$\sum_i h_i = 0 \tag{3}$$

for an ordinary or a hedge portfolio respectively. After rescaling the quantities h_i satisfying (1) may also be taken to satisfy either (2) or (3). Thus it is evident that apparently risk free portfolios will exist if $N > T$. In fact due to measurement error in the cells of the covariance matrix it can even happen that for some n

$$h^t \Omega^{hist} \ h \ < \ 0$$

i.e. that some portfolios have apparently negative risk. This circumstance causes serious difficulties if the investor builds his portfolios using a mean-variance optimization technique. For the optimizer will mistakenly believe that by mixing a negative risk portfolio into a regular portfolio it can reduce the risk of the

resulting portfolio. Thus the estimated risk of an optimized portfolio may be subject to serious downward bias if the covariance matrix is not positive definite.

The limitations of the historical covariance matrix motivate the search for a better estimator of the asset covariances. The standard solution is to invoke a factor model. A factor model is a linear model for asset returns such that

$$r_i(t) = \sum_{j=1}^{m} X_{ij}(t)\ f_j(t)\ + \epsilon_i(t)$$

where $X_{ij}(t)$ is termed the exposure of asset i to factor j, $f_j(t)$ is termed the return to factor j and $\epsilon_i(t)$ is termed the specific return to asset i. It is further assumed that the factors f_j capture all common sources of return between assets, or equivalently that

$$\mathrm{cov}(f_j(t),\ \epsilon_i(t)) = 0 \tag{4}$$

for all factors $f_j(t)$ and specific returns $\epsilon_i(t)$ and that

$$\mathrm{cov}(\epsilon_i(t),\ \epsilon_k(t)) = 0 \tag{5}$$

for distinct assets i and k. In this case with a bit of algebraic manipulation one can show

$$\Omega = XFX^t + \Delta$$

where

$$F_{ij}(t) = \mathrm{cov}(f_i(t),\ f_j(t))$$

and

$$\begin{aligned} \Delta_{ij}(t) &= \mathrm{cov}(\epsilon_i(t),\ \epsilon_j(t)) && (6) \\ &= \begin{cases} \mathrm{var}(\epsilon_i) & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} && (7) \end{aligned}$$

Suppose there are M factors. Then the F matrix requires us to estimate $M(M+1)/2$ distinct covariances and the Δ matrix requires us to estimate N variances. For M much smaller than N, we see that many fewer covariances need to be estimated for a factor model than for a historical covariance matrix. Thus, F and Δ may be estimated historically, i.e.

$$F_{ij}^{hist}(t) = \text{cov}(\{f_i(u)\}_{u=1}^{t}, \{f_j(v)\}_{v=1}^{t}) \quad (8)$$
$$\Delta_{ij}^{hist}(t) = \text{var}(\{\epsilon_i(u)\}_{u=1}^{t}) \quad (9)$$

within reasonable error. Furthermore if $M < T$ then

$$\Omega^{fact} = X\ \hat{F}^{hist}\ X^t + \Delta^{hist}$$

will be positive definite.

There are two basic approaches for generating factor models, which are termed the exploratory and confirmatory approaches respectively. The exploratory approach assumes that the returns are generated by a factor model but that nothing is known about the factor model

$$r = Xf + \epsilon$$

Various statistical techniques can be then be applied to simultaneously estimate X_i and $f(t)$ from the data $\{r_i(t)\}$. In this method X captures all the cross-sectional variation in i and $f(t)$ captures all the temporal variation. Even so, however, X and $f(t)$ are not uniquely defined, but rather are determined only up to a rotation of the exposures. Thus the factors extracted by this technique are not directly interpretable. Interpretability of factors is an important consideration if a risk model is to be used for active portfolio management. In active management risks are deliberately taken in the effort to earn compensatory return. Thus judgments must be formed as to whether or not one is willing to take on risk along a particular dimension. If the dimension is a statistical construct without an economic interpretation, there is little basis from which to form such judgments. It is usual in exploratory factor analysis to apply a rotation to the extracted factors in the hopes of arriving at an interpretation of the factors. The value of this interpretation, however, relies entirely on the analyst's judgment.

The confirmatory factor approach assumes that *a priori* information is available about the factor structure. In the returns based approach one assumes that the factor returns $f_j(t)$ are known. Then the exposures $X_{ij}(t)$ are found by regressing the asset returns on the factor returns. For instance the Capital Asset Pricing Model (CAPM) assumes a single factor return, namely the market return $m(t)$ in excess of the risk free rate $r_0(t)$ and it determines an asset exposure, the historical beta β^{hist} by regressing asset returns in excess of the risk free rate on the market excess return

$$r_i(t) - r_0(t) = \alpha + \beta_i^{hist}\ [m(t) - r_0(t)]\ + \epsilon_i(t)$$

A limitation of the returns based approach is that the estimated factor exposures may not be very interpretable. A variant, known as style analysis, attempts

to correct this defect by carrying out a least squares estimation in which the exposures are restricted to lie in an *a priori* reasonable range. For instance restricting the exposures to lie between 0 and 1 and to sum to 1 allows them to be interpreted as weights which describe how the factor returns are mixed together to best approximate the asset return. Note that if the data does not conform to the imposed restrictions then in general

$$\operatorname{cov}(f_i(t),\ \epsilon_j(t)) = 0$$

will fail to hold, so style analysis cannot simultaneously guarantee interpretable exposures and a consistent factor structure.

In contrast to the returns based analysis, the exposure based approach to confirmatory factor analysis assumes that the exposure matrix X is known *a priori*. The factor returns f(t) are then estimated by regressing the asset returns on the exposures. The exposure approach differs from other factor modeling methods in that

1) the interpretive structure is unambiguous

2) the exposure matrix can vary dynamically through time.

However one generates a factor structure, one is faced with the problem of assessing its adequacy. For exploratory techniques one is guaranteed to find a structure which meets the basic assumption (4) and (5) of the factor model over the time period in which the model is estimated. The essential assessment then is whether the factor exposures estimated in this way remain stable in subsequent time periods. For returns based confirmatory factor analysis stability of the factor exposures is again the basic criteria of success. For exposure based confirmatory analysis by contrast the technique only guarantees that property (4) will hold. Thus the essential test of the model is verifying how well property (5) holds – i.e. are the specific returns of distinct assets of zero covariance within measurement error?

All confirmatory factor analysis techniques assume that some prior information is available about asset returns. Here it is worthwhile to note the connection between the analysis of prices (i.e. security valuation) and the analysis of returns. We illustrate this connection by a simple case. The dividend discount model relates the price P_i of asset i to its dividend rate D_i and a suitable discount rate r_{ddm} by

$$P_i(t) = \frac{D_i(t)}{r_{ddm}(t)}$$

or equivalently

$$\ln P_i(t) = \ln\ D_i(t) - \ln\ r_{ddm}(t)$$

Let us assume that $D_i(t)$ is so slowly varying we may consider it constant. Then the derivative of log price with respect to time is just the return, so

$$r_i(t) = \frac{d}{dt}\ \ln\ P_i(t) = -\frac{1}{r_{ddm}}\ \dot{r}_{ddm}(t)$$

The terms on the right hand side of this equation are independent of the asset identity i. Hence this pricing model asserts that all asset returns will be governed by a common factor $\dot{r}_{ddm}(t)$. Recognizing that the pricing model ignores relevant data, a more reasonable hypothesis would be

$$r_i(t) = \beta\ \dot{r}_{ddm}(t)\ +\ \epsilon_i(t)$$

where $\epsilon_i(t)$ captures the effect of detail not included in the pricing model. Taking the mean over assets i, one would expect the effect of the $\epsilon_i(t)$ to average out, so the mean return should be

$$m(t) = \beta\ \dot{r}_{ddm}(t)$$

and on substitution

$$r_i(t) = m(t) + \epsilon_i(t)$$

This hypothesis can then be checked by performing the regression

$$r_i(t) = \alpha_i + \beta_i m(t) + \epsilon_i(t)$$

which is in essence the capital asset pricing model. Thus one sees that economic insights we may have into asset valuation can be translated into insight into asset returns, and of course conversely.

Similarly any intuitions we may have as to the sources of financial risk may be used to generate factor concepts which can be tested for their cross-sectional power in explaining returns. For instance it is intuitive that the risk of a security should be related to its position in the issuer's capital structure.

One important insight which follows from valuation theory concerns the composition of the universe of assets for which we attempt to construct factor models. Securities are typically valued by comparison with similar securities. It follows that we should attempt to apply factor analysis to a reasonably homogeneous collection of assets, and not to any miscellaneous assortment of securities.

An interesting aspect of the confirmatory factor approach is that it is possible to specify ambiguous models. For instance, we might propose that the return for an asset has two common factors:

1) a world equity return factor
2) a country equity return factor

Thus we would write

$$r_i(t) = \alpha_i + X_{i0}\, f_0(t) + \sum_{j=1}^{m} X_{ij}\, f_j(t) + \epsilon_i(t) \tag{10}$$

where

$f_0(t)$	is the global equity return factor
X_{i0}	is the exposure of asset i to $f_0(t)$
$f_j(t)$	is the return to the j^{th} country return factor
X_{ij}	is the exposure of asset i to the j^{th} country and in particular $X_{ij} = 0$ if asset i is not related to country j
$\epsilon_i(t)$	is the specific return of asset i

As written this model could represent either a returns based analysis or an exposure based analysis. In either case the model is ambiguous. Consider first the exposure based case. In the simplest formulation $X_{ij} = 1$ if asset i is domiciled in country j and otherwise $X_{ij} = 0$. Let $M_0(t)$ be the mean return over all assets at time t. Similarly let $M_j(t)$ be the mean return over all assets in country j at time t. If we attempt to fit the model (10) by least squares regression we will find that one solution is

$$f_0(t) = M_0(t)$$

$$f_j(t) = M_j(t)$$

However for any constant c we shall find that

$$f_0(t) = M_0(t) + c$$

$$f_j(t) = M_j(t) - c$$

is an equally good solution, for when this solution is substituted into (10) the c terms on the right hand side will cancel out. The difficulty is simply that

every asset which is in the world is also in a country and thus we can not fully disentangle the world effect from the country effect. There are two approaches to resolving this difficulty:

1) we can impose a restriction on our model which removes the ambiguity, essentially by making a particular choice of the constant c. An obvious choice is $c = 0$. This amounts to forcing

$$\sum_j f_j(t) = 0$$

so that the world factor captures the mean equity return and the country factors capture the deviations around that return attributable to country exposure. In this case we say that the country factor returns are *active* to the world factor return.

2) Alternately we can drop the world factor from the regression and fit the model with only country factor exposures. The country factor returns are in this case total factor returns. If we wish we may then define the world factor return as the mean of the country factor returns and define the active country factor return as the differences

$$f_j(t) - f_0(t)$$

It will be seen that there is no essential difference in these two approaches. In each case an ambiguity inherent in the data is removed by imposing an interpretation and as the interpretation is the same in both cases the actual numbers calculated will agree.

Although we have illustrated model ambiguity in the exposure based approach, the same difficulty arises in the returns based approach. For suppose the factor returns $f_j(t)$ are measured by various index portfolio returns. Then in a consistent measurement framework $f_0(t)$ will be the mean of the $f_j(t)$ and so the regression equation (10) will be singular as before. In this case the ambiguity will arise in the definition of the exposures and an interpretive structure will have to be imposed as before to eliminate the ambiguity.

2.2 Global Equity Factor Models

Having reviewed the basic aspects of factor models, we turn now to a survey of the published literature in this field. We admit immediately that our survey has no pretense of being comprehensive. Cognoscenti will also recognize that

approximately half the cited literature is due to BARRA researchers. Judging by the bibliographies of non-BARRA researchers, however, we do not believe that this predominance is due to a home bias in sampling the literature. We would like, however, to apologize for any oversights in reviewing this literature. As the global factor modeling literature has grown slowly over time, a chronological survey provides an attractive route into the material.

As a historical backdrop it is useful to recall certain facts. Foreign investment by U.S. investors declined sharply in the 1930's as the world depression uncovered problems in foreign government bond flotations and weak financial controls in holding company structures (e.g. the Swedish March scandal). In the forties war strengthened this tendency, while in the fifties investment opportunities in the U.S. market gave little reason to search abroad. The sixties brought changes, however. On the one hand the postwar economic recoveries in Europe and Japan were hailed as miracles. On the other hand the formulation of the Capital Asset Pricing Model stimulated thinking about risk and diversification. U.S. based investors began to invest abroad again. This trend met with a temporary setback due to the scandals associated with Investors Overseas Services, which collapsed in 1971. The collapse of the more respectable nifty-fifty stock market in 1974, however, emphasized that buying blue chips was not risk free.

Besides a market collapse, 1974 also brought three seminal papers on risk control. Solnik investigated the diversification benefits of international investing by building portfolios ["Why not Diversify Internationally Rather than Domestically?", Financial Analysts Journal, July-Aug 1974, p. 48–54.] He found that diversifying by country produced greater diversification benefits than diversifying by industry within a country. Lessard applied returns based analysis to market portfolios to probe into the sources of international return ["World, National and Industry Factors in Equity Returns", The Journal of Finance, 1974, p. 379–391.] He found that the CAPM, i.e. a single factor model based on a capitalization weighted market return, extrapolated poorly to the international setting. The basic difficulty was that in this study period the U.S. and Canada comprised more than 50% of world cap and so the attempt to formulate a cap weighted world factor resulted in a poor fit for the rest of the world. Accordingly, Lessard was led to a multi-factor model in which the world factor was represented by an equal weighted average of market returns. Country and industry factors were identified as additional sources of return variation. The single factor CAPM approach was challenged in a single market setting by Rosenberg and Marathe ["Extra Market Sources of Covariance in Security Markets", Journal of Financial and Quantitative Analysis, March 1974, p. 263–274.] They developed an exposure based multi-factor model of U.S. equity returns. Going beyond academic work, Rosenberg transformed this model into a commercial risk measurement service, today known as BARRA.

In 1976 Lessard further refined his returns based approach to international equity risk ["World, Country and Industry Relationships in Equity Returns: Implications for Risk Reduction through International Diversification", Financial Analysts Journal, Jan-Feb 1976, p. 32–38.] In this analysis he worked at the asset level, as opposed to the market portfolio level of his earlier paper. He found evidence of a world effect which explained on average 19% of variation in equity return, and a country effect which explained an additional 22% of equity return. By contrast, industries explained only 6% in addition to the world effect. Although, Lessard's analysis focused on diversification benefits, his findings were favorable to a school of active management which favored top-down country selection and which tended to hold the local market portfolio in each country.

Following in Rosenberg's footsteps, Grinold, Rudd and Stefek developed a risk model of international equity returns in 1989 ["Global Factors: Fact or Fiction", Institutional Investor, Fall 1989] This paper presented what is today the core of the Global Equity Model (GEM). In this analysis return to a stock is decomposed into a currency part and a local market part

$$r_{i,n}(t) = r_{i,c}(t) + r_{i,l}(t) + r_{i,c}(t)\ r_{i,l}(t)$$

where

$r_{i,n}(t) =$ return to asset i over period t in numeraire currency.

$r_{i,c}(t) =$ return to the local currency of asset i in numeraire currency over period t plus the risk free return to the local currency of asset i over period t

$r_{i,l}(t) =$ the excess return to asset i in its local currency over period t, where the excess return is the total return less the risk free rate, both measured in local currency.

This decomposition is then linearized by dropping the cross product term which in most cases is a term of negligible magnitude[1].

The local excess return is then modeled as

$$r_{i,l}(t) \;=\; \beta_i(t)\ f^1_{d(i)}(t) \;+\; f^2_{k(i)} \;+\; \sum_{j=1}^{4} X_{ij}(t)\ f^3_j(t) \;+ \epsilon_i(t)$$

where

[1]The cross product term is most significant for volatile currencies and markets where the currency and market are correlated. This circumstance is more usual in the emerging markets than in the developed markets. Even so the error induced by dropping the cross product term is generally acceptable.

$\beta_i(t)$	is the historical beta of asset i over the five year period $[t-60, t)$
$d(i)$	is the country of domicile of asset i
$k(i)$	is the principle industry of asset i as of the start of period t
$f^1_{d(i)}(t)$	is the estimated return to $d_{(i)}$ country $d_{(i)}$ in period t
$X_{ij}(t)$	are four statistical characterizations of asset i as of the start of period t

Traditionally

j=1	is termed VOLATILITY
j=2	is termed SUCCESS
j=3	is termed SIZE
j=4	is termed VALUE

$f^3_j(t)$	is the estimated return to the jth factor over period t
$\epsilon_i(t)$	is the specific return to asset i over period t.

Here the quantities $f^1_d(t), f^2_k(t), f^3_j(t)$ and $\epsilon_i(t)$ are estimated by a cap weighted regression of asset returns against structural factors.

Grinold, Rudd and Stefek applied this model to generate various insights into the world equity market:

1) They showed that country returns as measured by market index returns are "colored" by the industry composition of the local market portfolio.

2) They found that the crash of 1987 was primarily captured by the industry factors, rather than the country factors.

3) While both industry and country factors are important contributors to volatility, countries accounted for somewhat more volatility. In general countries explained 27% of the variance and industries explained an additional 8%. This finding can be interpreted as indicating that country timing is a more fruitful source of alpha than industry timing.

4) They proposed the correlation between the local market index and local country factor as a measure of how integrated or segmented the local market is within the global economy, with a higher correlation indicating greater segmentation. This measure produces intuitive results.

In 1992 similar matters were examined by Richard Roll ["Industrial Structure and the Comparative Behavior of International Stock Market Indices", Journal of Finance, March 1992 p. 3-41.] Roll's focus was explaining the differences in

behavior between stock market indices. Roll's study was somewhat handicapped by limited access to asset level information. However, he was able to explain the behavior of indices by:

1) differential diversification by name
2) differential diversification by industry
3) differential industry composition
4) exchange rates

Thus his results were broadly consistent with the finding of earlier authors which emphasized the importance of industry effects in understanding market index returns.

Also in 1992 Beckers, Grinold, Rudd and Stefek examined the behavior of European equity returns ["The Relative Importance of Common Factors Across the European Equity Market", Journal of Banking & Finance (16), 1992 p. 75–95.] Their model was essentially a restriction of their 1989 model to the European countries. A primary motive for the study was to see if a regional equity market exhibited similar or different behavior from the global market. The study showed the European market to exhibit much the same behavior as the global market. This finding was somewhat surprising to the authors, who had thought that the economic integration fostered by the European community would reduce the importance of country factors relative to industry factors. However, the evidence adduced did not document such an effect.

In 1992 Drummen and Zimmerman also made an analysis of European equity returns ["The Structure of European Stock Returns" Financial Analysis Journal, July-Aug 1992, p. 15–26.] They applied an exploratory factor analysis method to extract factors from asset returns. They were able to interpret these factors as largely country effects with some additional industry effects. They also investigated the data using returns based confirmatory factor analysis. The factor returns are either index returns or the results of orthogonalizing index returns to other return series. The factors utilized were world, European and national market returns, as well as two currency returns (US-to-local and ECU-to-local) and industry returns. They found the national market factors to be dominant with industry factors being of secondary importance and currency factors of tertiary importance. The effect of the world and European factor was largely transmitted through the country factor.

In 1993 Connor and Korajczyk introduced a new statistic, denoted EP, for measuring the explanatory power of factors in a model ["A Test for the Number of Factors in an Approximate Factor Model", The Journal of Finance, Vol. 48 no. 4, Sept 1993, p. 1263–1291.] A commonly used method for measuring the

importance of a factor is to fit the model both including and excluding the factor. The model fit in each period is assessed by the R^2 statistic and the model's overall performance is assessed by the time-series mean of R^2. Comparing the mean R^2 for the model with and without the factor gives a measure of the factor's marginal contribution to fit. Connor and Korajcyzk introduced EP as a substitute for the time series mean R^2. Then the factor's marginal contribution to explanatory power is measured by the difference in EP for the model with and without the factor. The advantage of this technique is that it can measure the explanatory power of the model's intercept term, whereas the R^2 method cannot.

In 1995 Heston and Rouwenhorst re-examined European equity returns ["Industry and Country Effects in International Stock Returns", Journal of Portfolio Management, Spring 1995, p. 53-58] using exposure based analysis. In their model equity returns are decomposed into
1) an equal weighted mean European return
2) an industrial factor effect
3) a country effect
4) a firm specific effect
The industrial and country effects are taken active to the mean return. In agreement with prior authors, they found country effects to dominate industry effects.

This work was extended by Becker, Connor and Curds in 1996 ["National Versus Global Influences on Equity Returns", Financial Analyst Journal, March-April 1996, p. 31-39.] Their model decomposed local excess returns into

1) a global market factor
2) a country factor
3) an industry factor
4) a firm specific effect

where the industry and country factor returns are made active to the global market factor. The authors considered the nested models produced by dropping various effects. To compare the different models they used the time-series mean adjusted R^2 and the time-series mean adjusted EP. The R^2 criteria is insensitive to the contribution of the global market factor, whereas the EP statistic captures the explanatory power of this term. The significance of the different factor is summarized as

Factor	R^2	EP
global market	–	21.1%
countries	17.6%	15.1%
global sectors	1.9%	1.6%
global industries	4.3%	3.5%
local industries	12.8%	10.8%

The authors also examined the shifts in relative importance of the different factors over time. For the world as a whole they found only slight evidence of trends. When they estimated the model over just 4 EU countries, however, they found country factors to be of declining importance and industries to be of increasing importance. In particular, the mean correlation between the country factors increased by 2.5 percentage points per year over the 8/83 - 8/93 period. This trend is statistically significant with a T statistic of 4.0. The authors interpreted this finding as evidence of increasing integration within the EU region in contrast to a static level of integration in the world as a whole.

Expanding the view considerably, Chaumeton, Connor and Curds considered fitting a factor model to global stock and bond markets ["A Global Stock and Bond Model", Financial Analysts Journal, Vol. 52 no. 6 1996, p.65–74.] These authors formulated two models. In both models the factors which enter the analysis are

1) shift - bond sensitivity to parallel yield curve shifts (i.e. modified direction)
2) twist - bond sensitivity to a yield curve twist
3) an equity market factor
4) an equity size factor
5) an equity value factor
6) an equity duration factor (essentially inverse yield.)

In one model these are global factors, while in the other model these are national factors. The global model explains 21% of stock volatility and 60% of bond volatility. The marginal explanatory power of the national factors is 13% for stocks and 35% for bonds. The two most significant factors are the equity market return and the bond market shift. As global factors the explanatory power of these factors is 19% for stocks and 59% for bonds. As national factors the explanatory power is 37% for stocks averaged across markets and 91% for bonds averaged across markets. These factors are also the most correlated. The correlation between the global factors is 35% and the mean correlation between a national equity market factor and the corresponding national shift factor is 41%. These results are of course entirely consistent with the interpretation of the equity market factor suggested by the dividend discount model.

In 1998 Eckhard Freimann further examined European integration ["Economic Integration and Country Allocation in Europe", Financial Analysts Journal, Sep-Oct 1998, p. 32-41.] He examined correlations in returns for country indices over various subperiods from January 1975 through December 1996. He found correlations to be increasing, with an intuitive and interpretable pattern of increase when examined in the context of various macroeconomic measures of monetary and economic integration. However, he also found that correlations had not yet converged to the levels which would apply in the absence of any country effects. Thus he inferred that country factors still effected European equity returns.

A different perspective was provided by Richard Weiss ["Global Sector Rotation: A New Look at an Old Idea", Financial Analysis Journal, May - June 1998, p. 6-8.] In a thought piece he argued against the shortsightedness of global equity alpha strategies based purely on country and currency selection. He pointed out that such strategies failed to exploit the opportunities implicit in timing industry factors and often generated unintended industry timing side bets. As a corrective he argued for explicit industry selection. Weiss saw increasing global economic integration as a threat to country and currency strategies. Thus he felt industry timing had a more promising future.

In 1999 also a new line of inquiry opened with a focus on companies with multinational operations. Chaumeton and Coldiron in particular examined whether these companies should be analyzed differently from purely domestic enterprises ["Global Companies - A New Asset Class?", Barclays Global Investors Equity Research Report, February 1999.] They defined a company to be a "global company" if more than 50% of its sales are earned outside its home market and it is in the top quartile of its local capitalization range. They found local markets ex global companies to exhibit somewhat greater internal homogeneity and somewhat less external correlation than the full local market. Thus global companies appear to play a role in connecting local markets. However, the evidence of this effect was weak.

In 1999 Rouwenhorst returned to an examination of integration in European equity markets ["European Equity Markets and the EMU", Financial Analyst Journal, May-June 1999, p. 57–64.] His model followed his earlier work in 1995 (with Heston) in decomposing returns into a global market factor, active country factors and active sectors. In a variation in technique the regression implementing the decomposition was now cap weighted. He assessed the importance of country and sector effects by looking at the mean absolute factor return in each month averaged over a 36 month moving window. He found both effects to be somewhat variable, with country effects being more variable. Consequently a perception of trend will be sensitive to the choice of study period. Rouwenhorst

concluded that country effects continued to dominate sector effects.

In 2000 Cavaglia, Brightman and Aleed re-examined the question of country factors dominating industry factors. ["On the Increasing Importance of Industry Factors: Implications for Global Portfolio Management", Financial Analysts Journal, Forthcoming.] Their model is

$$r_{ij}(t) = d(t) + f_i(t) + g_j(t) + \epsilon_{ij}(t)$$

where

$r_{ij}(t)$	is the return to the ith industry portfolio in country j excess of the local risk free rate
$d(t)$	is the global market return
$f_i(t)$	is the ith global industry factor
$g_i(t)$	is the jth country factor
$\epsilon_{ij}(t)$	is the residual

The model is estimated as a cap weighted regression with constraints on the global industry and country returns so that they are active to the global market return. Unlike earlier work, this analysis is based on weekly returns rather than monthly returns. As a result higher discrimination of secular behavior is achieved. The authors summarized the importance of country industry factors using the mean absolute return measure employed by Rouwenhorst. They find that the importance of industries relative to countries has been trending upwards since November 1994 and that industries have dominated countries since November 1998. They point out that given the importance of industry effects, a home biased benchmark may lead to very inefficient asset allocation if the home market is significantly underweight important global industries.

2.3 Prior Barra Research

In addition to the published literature, we may also take note of BARRA's internal research, much of which has previously been published in the BARRA research seminar series.

BARRA's basic approach to global equity factor modeling is closely based on the Grinold, Rudd, and Stefek model of 1989. In this approach local market residual returns

$$\tilde{r}_i(t) = r_i(t) - rf_i(t) - \beta_i(t)\, m_i(t)$$

are calculated, where

$r_i(t)$ is the return to asset i in local currency
$rf_i(t)$ is the risk free rate for the local currency of asset i
$\beta_i(t)$ is the beta from a five year CAPM regression
$m_i(t)$ is the local cap weighted equity index return

The market residual returns are then fit to an exposure based factor model of the form

$$\tilde{r}_i(t) = \sum_{j=1}^{N_i} X_{ij}(t) f_j^1(t) \; + \sum_{k=1}^{4} X_{ik}(t)\; f_k^2(t) \; + \epsilon_i(t)$$

Here the $X_{ij}(t)$ are N_i industry exposures where for every asset i there is a unique j such that $X_{ij}(t) = 1$ and for all $k \neq j$ $X_{ik}(t) = 0$. The quantities $f_j^1(t)$ are interpreted as returns to globally defined industries. The $X_{ik}(t)$ are four statistical characterizations of asset i, termed SIZE, SUCCESS, VALUE and VARIABILITY IN MARKETS. These measures are based on asset characteristics normalized against the local market. By contrast the returns $f_k^2(t)$ to these factors are estimated globally. The estimation of this model is carried out over approximately 2000 assets, drawn from some 25 so called developed markets [2]. The commercial version of this model is known as GEM, an anacronym for Global Equity Model. It exists in two different versions distinguished by modestly different industry classification schemes and estimation universes.

The principal difference between GEM and the 1989 model is that in GEM the market return $m_i(t)$ is calculated exogenously, whereas in the 1989 it was included in the common factor regression. Recall that in the 1989 paper the market exposure was set equal to the historical betas $\beta_i(t)$. Empirically it was found that the vector $\beta(t)$ was sufficiently close to the unit exposure vector as to induce significant multi-collinearity with the industry exposure matrix, thus leading to an unstable regression. The shift to an exogenously determined market return removed this difficulty. In essence a problem of model identification was removed by establishing a priority order between country and industry effects which gave first weight to countries.

Since its initial formulation GEM has been the subject of a vigorous program of internal criticism and research. This program has both affirmed important aspects of the model, while also uncovering limitations and indicating directions for further model development.

[2]the exact composition of the estimation universe varies through time.

In 1991 Stefak examined the possibility of using regional sectors instead of global industries ["World Equities: Regional Sectors or Global Industries", International Portfolio Management Seminar 1991, Section P.] He found the following adjusted R^2:

Global Industries	32%
Regional Sectors	34%
Regional Industries	40%

The behavior of the sectors in different regions generally differed. Much of this difference derived from Japan, as Europe and North America generally exhibited more similar behavior. Thus, these findings tended to throw doubts on global industries as a concept.

In 1992 Grinold and Drach examined the opportunity to improve the modeling of specific risk in GEM ["International Specific Risk", 1992 Equity Research Seminar, Section N.] In the original GEM model the specific risk vector is modeled as

$$\epsilon(t) = N(0, \Delta(t))$$

where $\Delta(t)$ is a diagonal matrix whose elements are estimated by a 60 month moving window, i.e.

$$\Delta_{ii}(t) = \text{var}\ (\{\epsilon_i(s)\}_{s=t-60}^{t})$$

In this approach each asset is treated independently of all others. Research showed that useful cross-sectional information could be derived to improve the model. In this approach the mean specific return is estimated for each market and the asset relative variation is found around that mean. Thus one writes

$$\Delta_{ii}(t) = S_{d(i)}(t)\ R_i(t)$$

where

$d(i)$ is the country of asset i
$S_{d(i)}(t)$ is the mean specific return for country $d(i)$
$R_i(t)$ is the market relative specific return for asset d

A forecast for $S_{d(i)}(t)$ was arrived at by a model of the form

$$\hat{S}_{d(i)}(t) = a_{d(i)} + b_{d(i)} S_{d(i)}(t-1)$$

Similarly $R_{d(i)}(t)$ was forecast by a model of the form

$$\hat{R}_i(t) = \sum_j c_j Y_{ij}(t)$$

where the variables Y_{ij} included industry effects, size, yield and volatility effects and lagged realizations of the absolute specific returns $|\epsilon_i(t)|$. Out of sample tests showed an improvement in forecasting ability for $\sqrt{\Delta_{ii}(t)}$ from an R^2 of 20% for the moving window model to one of 31% for the new model. An implication of this analysis is that even in the context of a global model there could be considerable value to considering an asset in relation to the assets in its local market.

The focus on local market information was developed in a different direction by Kahn and Torre ["GARCH Scaling of the Covariance Matrix", unpublished, 1992.] GARCH is a time series technique developed by Engle ["Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation", Econometrica, Vol. 50 no. 4, July 1982, p.987-1002.] In this technique a time series $x(t)$ is modeled as

$$x(t) = N(\hat{x}(t),\ \sigma^2(t))$$

where $\hat{x}(t)$ is a model for $x(t)$ and $\sigma^2(t)$ is modeled by standard time series methods. For instance a GARCH(1,1) models $\sigma^2(t)$ as

$$\sigma^2(t) = w + \alpha\, \sigma^2(t-1) + \beta\, \varepsilon(t-1)^2$$

where $\varepsilon(t) = x(t) - \hat{x}(t)$ and w, α and β are constant parameters. Kahn and Torre showed that the equity market return as measured by a broad market index could be modeled by a suitable GARCH model to arrive at a forecast of total market volatility $\sigma^2(t)$. This forecast could then be used to rescale a factor covariance matrix to arrive at a more accurate risk forecast. Building on this insight Hui systematically examined different implementations of this concept in local markets worldwide and the implications for scaling the GEM covariance matrix. His conclusion was reported in 1997 ["New Approaches to Global Equity Risk", 1997 Equity Research Seminar, Section H.] He wrote the elements of the factor covariance matrix F in the form

$$F_{ij} = \sigma_i \, \sigma_j \, \rho_{ij}$$

where the quantities σ_i capture the factor variances and the quantities ρ_{ij} capture the factor correlations. The volatilities were estimated by either univariate GARCH models or by exponential weighted moving average (EWMA) models. The correlations were also estimated by an EWMA model. These changes significantly improved the F matrix over its original estimate by means of an expanding window sample covariance. In focusing on market factors in particular, Hui found that the times series properties of local markets differed around the world and that it was best not to impose a common global structure.

Hui also examined possible changes to the exposure matrix. Since the country factor returns are the most important group of factors for explaining variance, the country exposures are a natural point to focus on for possible refinement. One observation is that many companies actually do business in multiple countries and so possibly one should allow assets to be exposed to more than one country. Although an engaging idea, Hui reported that implementable versions of this concept had only marginal modeling impact ["Multiple-country Allocation of International Stocks", 1994 Equity Research Seminar, Section R.] Essentially one needs a method to assign country exposures in the X matrix. One approach would base these exposures on reported geographic distributions of sales, assets or earnings. Unfortunately company reporting of this data is quite spotty and unsystematic. Applying this approach results in allocating only 2% of world cap outside the country of domicile. The impact of this change on the model performance is negligible. As a way around this difficulty, one can attempt to estimate country exposures using style analysis. This analysis starts with return series $m_j(t)$ and finds weights X_{ij} such that

$$\sum_t \left[r_i(t) \; - \; \sum_j X_{ij} \, m_j(t) \right]^2$$

is minimized subject to the requirement

$$\sum_j X_{ij} = 1$$

One could take the $m_j(t)$ to be country index returns. However it is found that the technique produces more intuitive results if one takes $m_1(t)$ to be the local market return for the country of domicile of asset i and for $m_j(t)$ $(j > 1)$ to be the portion of the return to market j which is uncorrelated with $m_1(t)$. When this technique is applied approximately 4% of world cap is allocated outside

the country of domicile. In one test moving from single country exposures to multi-country exposures raised the mean model R^2 from 35% to 36%.

While the research on multi-country exposures appears to be a marginal result, it actually has an interesting interpretation. The style analysis result shows that companies are strongly related to their local market factor. As we saw earlier one interpretation of the market return is that it represents fluctuations in the discount rate applied to pricing equities. Accepting this interpretation of the market factor the implication of the style analysis is that most companies are priced in their local market. Even if the cash flows are generated around the world, they are discounted locally and this discounting is largely independent of where the cash flows are generated. Can we accept the hypothesis that companies are priced locally? For capital surplus countries it seems logical that the marginal investor should be domestic, and thus in those countries one would expect the companies to be priced locally. For capital deficit countries, however, it seems that the marginal investor should be an international investor and so companies would be priced in the world market place. Even so, however, international investors may only distinguish between the domestic focused and multi-national exposed enterprises of such countries as a second order effect. Thus even in capital deficit countries it might be that the companies were being priced as a group even if the marginal investor is not locally based. Follow up studies which distinguish between capital surplus and deficit countries might throw further light on this point.

The attempt to estimate multi-country exposures highlighted that one of the imitations of exposure based factor analysis in the global setting is the comparative absence of detailed fundamental data which is comparable from country to country. The style analysis approach to determining country exposures indicates a possible solution to this problem. More generally one could ask if a returns based analysis might not be more effective in a global setting than the exposure based analysis. To investigate this point Hui fit a returns based model in which asset returns were attributed to country, currency and industrial sector returns. A particular orthogonalization procedure was applied to the factor return series to remove the ambiguities in exposure loadings that would otherwise result from correlations in the factor returns. Hui assessed the model based on an examination of the factor loadings ["Approaches to Global Equity Risk", 1996 Equity Research Seminar, Section P.] He found that the results were frequently unintuitive, for instance about 60% of the assets were not exposed to their local market. Factor loadings were very unstable, with the regression of loadings estimated in the 1989-95 period on the loadings estimated in the 1983-88 period possessing an R^2 of only 1%. Finally only about 60% of the assets possessed adequate data for the methodology to be applied. Based on these and other findings Hui concluded that a returns based analysis was distinctly inferior to

an exposure based analysis.

In searching for ways to improve GEM a natural approach is to examine the performance of portfolio risk forecasts. For portfolios broadly diversified by country we find that GEM does a reasonably good job. For portfolios concentrated in a single country, however, GEM's performance is less satisfactory. Such portfolios may also be analyzed using a model of the local equity market. We generally find that the local model's performance dominates that of GEM, sometimes quite significantly. An interesting question which this observation raises is how well GEM can be expected to perform on portfolios which are concentrated in a single region (e.g. Europe or Latin America.) For portfolios drawn from regions located within the developed world we find GEM's performance to be reasonable. For portfolios located in emerging market regions, however, the performance is less satisfactory. Since the GEM estimation is dominated by assets in developed markets this result is not too surprising. However it suggests that factors may actually be regional rather than global. Thus, these observations motivate us to examine alternatives to the GEM factor structure.

Hui's first examination of the factor structure ["New Approaches to Global Equity Risk", 1995 Equity Research Seminar, Section P] examines two alternatives to GEM. The comparison of model structures was made on the basis of mean model R^2. The base case is GEM, which in the study period exhibited the following characteristics:

country factors	26.8%
global industries	20.1%
country & global industries	34.6%
countries & global industries & risk indices	36.2%

From this data we see that the marginal power of the industries is 7.8% and of the risk indices 1.6%. Although this comparison makes the risk indices seem relatively unimportant, it should be noted that there are only four risk indices, while there are 36 global industries. Thus the marginal power per industry is 0.22% while the marginal power per risk index is 0.4%. On this basis a risk index factor carries twice the power of an industry factor.

As an alternative to GEM Hui constructed a model in which countries are classified into seven regions and industries into seven sectors. For each combination of region and sector a factor is introduced. Thus for instance, one considers the North American Energy sector or the European Capital Goods sector. Estimating a model with countries and regional sectors results in an R^2 of 37%. The R^2 adjusted for the number of factors in this model is 35% as compared to 32.8% for the submodel of GEM estimated with only countries and global industries.

The marginal explanatory power of the regional sectors is 8.2% and the marginal power per factor is 0.17%. Examining the regional sector groupings one asks both whether these concepts appear meaningful and whether the correlations between the same sectors in different regions suggest that the regional sectors are actually subsections of global sectors. Here the evidence appears mixed. Japanese, North American and European sectors often appear to be meaningful concepts. However Asian sectors generally do not. Correlations between regional sectors suggest a global energy sector and possibly global transportation and basic industry sectors. Most sectors, however, do not show strong evidence of globalization. The basic trade-off in this factor structure as compared with GEM is to allow greater geographic specification with less industry specification. On balance the evidence indicates that this is a worthwhile trade-off.

The second alternative to GEM moved even further in the direction of geographic localization. In this analysis 23 countries were selected and in each market a factor model consisting of seven sector and four risk indices was fit. The adjusted R^2 of these models varied between 30% and 49% with a mean of 40%. To the country sector and risk index factors a currency factor was added to result in 12 factors per country or 276 factors in total. Principal components analysis was applied to this collection of factor returns. This analysis indicated that the first 100 principal components captured 99.4% of the variance. Examination of the first principal component shows that it captures 34% of the variance and that it may reasonably be interpreted as a world market factor. The second principal component suggests an interpretation as a continental Europe factor. Higher principal components are not easily interpretable. Examination of correlations suggests two geographic regions: Malaysia-Singapore and US-Canada. There is also evidence of a global energy sector. This analysis shows the benefits of country level analysis, while also affirming the existence of regional and global effects.

Hui further extended the country level approach ["Approaches to Global Equity Risk", 1996 Equity Research Seminar, Section P.] First he applied a number of analytic techniques to determine an efficient mapping of industries into economic sectors. Then he defined a standardized single country factor model based on 8 sectors and five risk indices. This model was estimated on each country separately resulting in some 700 factors. Hui regarded a 700 x 700 factor covariance matrix as unmanageable and so he explored two methods for reducing its dimensionality: principal components and imposition of a prior structure. The principal components method reduces the dimensionality to approximately 100, but the resulting factors are not very interpretable. As an example of a possible prior structure, one could impose the constraint that sectors in different countries are uncorrelated, thus Japanese Capital Goods would be forced to have a zero correlation with Chinese Basic Materials. A prior structure can reduce

the dimensionality of the covariance matrix, but the issue of what structure to impose is a significant one.

Hui continued to explore the approach to global analysis through single country analysis ["New Approaches to Global Equity Risk", 1997 Equity Research Seminar Section H.] In developing the standardized single country model he found that a single factor structure for all countries was not going to be successful. For countries with many assets, e.g. the US, a sector level analysis might be too coarse. By contrast, estimating even 8 sectors might be too fine a division in smaller markets with less than a 100 assets. To address this problem he developed a nested collection of standard factor structures so that the resolution of the model could be matched to the size of the capital market. How to combine the single country models into a global analysis remained an unsolved problem, however.

A different line of inquiry was opened by Chandrashekaran ["New Ideas in Risk Forecasting", 1998 Equity Research Seminar, Section H] and continued by Chandrashekaran, Hui and Rudd ["Narrow Markets and Global Stocks", 1999 Equity Research Seminar, Section F.] This inquiry began with two observations:

1. For many companies active around the world the impact of the economy of the country of domicile on the earnings of the company can be fairly minimal.

2. In the 1997-99 period many large companies had displayed differential market performance from smaller companies.

These observations lead to the hypothesis that the world equity market is segmented into a group of global companies which are priced globally and a collection of purely local markets. Any one national market would be a combination of global and local companies. One might look for the discount rate of the global segment to be a factor which would separate the performance of global companies within a national market from the purely local companies in that market. This global company factor might explain the differential performance of large and small companies that is observed if company size is correlated with being a global company. One would expect the global companies to be priced globally, the local companies to be priced locally and companies within a national market to be priced relative to one another. Thus the global companies would play a role in transmitting pricing between otherwise segmented local markets. The mix of local and global companies in a particular market might explain the relative correlation between that market and the world market, i.e. it would explain how closely the local pricing structure is tied to the global structure. This hypothesis is essentially the same as that investigated by Chaumeton and Coldiron. Chandrashekaran, Hui and Rudd investigated the hypothesis from a

number of angles. They found the problem of characterizing a global company to be a subtle one. From a returns perspective disentangling a global company effect from a size effect and global industry effects proves problematic. Their assessment was that global companies clearly exist but a return factor associated to being a global company is a yet unproven hypothesis.

A different approach to segmentation was taken by Conner and Herbert ["Regional Modeling of Western Europe Equities", Internal Report, 1998.] They investigated the factor structure of European Equities. They found this group of assets to divide into UK based assets and continental European assets. They found country effects which were diminishing through time. They fit a factor model based on

1) 19 UK industry exposures
2) 19 continental industry exposures
3) 16 country exposures
4) 6 risk indices (size, momentum, value, volatility, yield and a blue-chip indicator)

The country exposures and industry exposures are both zero-one variables, so to identify the model the countries are taken active to the industries. The performance of this model was compared against that of GEM by Curds and Herbert ["An Investigation into the Performance of the European Equity Model", Internal Report, 1999.] They found the European model to be superior to GEM in risk forecasting for European portfolios. Whether this superiority was due to a broader estimation universe or to differences in the factor structure was left unresolved, however.

2.4 Conclusions

Since the original formulation of GEM we have learned much about global equities:

1) The power of exposure based factor analysis has been confirmed. Exploratory factor analysis and returns based analysis have generally confirmed findings first achieved with GEM without generating fundamentally new insights themselves (Drummen and Zimmerman 1992, Heston and Rouwenhorst 1995.) At the same time the limitations of returns based analysis have been convincingly documented (Hui 1996.)

2) All studies have found country factors to be important. However the most recent studies find country factors to be of declining importance in Europe

(Becker, Connor and Curds 1996, Freimann 1998, Connor and Herbert 1998.) In discussions about the relative importance of country and global industry factors it is accordingly useful to distinguish the European situation from the rest of the world.

3) GEM found global industries to be important explanatory concepts. However, other studies suggest that this factor structure may be partially an analyst construct rather than an organic feature of asset returns. From a principal component analysis only the energy sector emerges as a clearly global industry (Hui 1995.) regional and country industries are found to carry substantially more explanatory power than global industries (Stefek 1991, Becker, Connor and Curds 1996.) Just within Europe, continental and UK industries are found to be separate concepts (Connor and Herbert 1998.)

4) The importance of different factors fluctuates through time (Rouwenhorst 1999.) Recently industry and size factors have been of considerable importance (Chandrashekaran, Hui and Rudd 1999, Cavaglia, Brightman and Aleed 2000.)

5) A number of studies have emphasized that global equities exhibit considerable inhomogeneities. Simple differences in available data handicap a number of studies (e.g. Chaumeton and Coldiron 1999, Chandrashekaran, Hui and Rudd 1999.) Specific risks are best understood in a local context (Grinold and Drach 1992.) Local markets exhibit distinct behaviors of market risk (Hui 1997.) Pricing appears to be local (Hui 1994.) Regional concepts appear to be valid in some but not all parts of the world (Hui 1996.) Local market factor structures differ (Hui 1997.) Attempts to define a subuniverse of truly global companies appear problematic (Chaumeton and Coldiron 1999, Chandrashekaran, Hui and Rudd 1999.) The UK appears to be only partially integrated into Europe (Connors and Herbert 1998.)

6) Risk indices are found to have high marginal explanatory power per factor (Hui 1995.) However they have been entirely ignored or relatively underutilized in most studies. Compared to single country factor models, fundamental (i.e. accounting) data are underutilized in the global setting. The principal difficulty is variations in disclosure and accounting standards around the globe.

Synthesizing all of this information we are led to a new vision of global equities. Whereas GEM saw global equities as a homogeneous group caught in a simple factor structure, we now see each local market as the homogeneous grouping with different markets linked together into a global matrix by various regional and global effects. The natural realization of this vision is to fit a factor model to each local market. The local models can be customized to each market to capture its special features and to best exploit the available data. The local analysis must

then somehow be integrated into a global analysis. The work of Hui has since 1995 been pointing in this direction. How to achieve the integration of local models has, however, been an elusive point. The fundamental contribution of this paper is the resolution of this difficulty, as described in the next section.

3 Construction of a Global Analysis

3.1 Combination of Models

Our goal is to integrate factor models of local markets into a global analysis. For notational simplicity let us first focus on the problem of integrating just two models. Specifically let (X^1, F^1, Δ^1) and (X^2, F^2, Δ^2) be two different factor models. In other words, if r^i is a vector of returns from market i then

$$r^i(t) = X^i(t)f^i(t) + \varepsilon^i(t)$$

where

$$\mathrm{cov}(f^i(t), f^i(t)) = F^i(t)$$

and

$$\mathrm{cov}(\varepsilon^i(t), \varepsilon^i(t)) = \Delta^i(t)$$

We may form new entities

$$\begin{aligned} r(t) &= \begin{pmatrix} r^1(t) \\ r^2(t) \end{pmatrix} \\ X(t) &= \begin{pmatrix} X^1(t) & 0 \\ 0 & X^2(t) \end{pmatrix} \end{aligned}$$

and fit the model

$$r(t) = X(t)f(t) + \varepsilon(t)$$

then

$$f(t) = \begin{pmatrix} f^1(t) \\ f^2(t) \end{pmatrix}$$

and

$$\varepsilon(t) = \begin{pmatrix} \varepsilon^1(t) \\ \varepsilon^2(t) \end{pmatrix}$$

If we assume the factor exposures capture all sources of common return between any two assets then the covariance matrix $\Delta(t) = \text{cov}(\varepsilon(t), \varepsilon(t))$ will continue to be diagonal and in fact

$$\Delta(t) = \begin{pmatrix} \Delta^1(t) & 0 \\ 0 & \Delta^2(t) \end{pmatrix}$$

The covariance matrix of common factor returns $F(t) = \text{cov}(f(t), f(t))$ will be

$$F(t) = \begin{pmatrix} F^1(t) & F^{12}(t) \\ F^{12}(t)^t & F^2(t) \end{pmatrix}$$

where $F^1(t)$ and $F^2(t)$ are as given in the local market models and

$$F^{12}(t) = \text{cov}(f^1(t), f^2(t))$$

is a new piece of data. Thus $(X(t), F(t), \Delta(t))$ will constitute a factor model for the union of the two local markets. In particular the asset-by-asset covariance matrix $\Omega(t)$ for the combination of the two markets is given by

$$\Omega(t) = X(t)F(t)X(t)^t + \Delta(t)$$

Given $\Omega(t)$ risk analysis may be performed on the union of the two markets exactly as it is performed on each market separately. Hence, in summary, we see that the only new piece of information required to join (X^1, F^1, Δ^1) and (X^2, F^2, Δ^2) into $(X(t), F(t), \Delta(t))$ is $F^{12}(t)$.

Let us consider how we might estimate $F^{12}(t)$. The simplest approach would be to form the sample covariance matrix

$$\hat{F}^{12}(t) = \text{cov}(\{f^1(u)\}_{u=1}^t, \{f^2(v)\}_{v=1}^t)$$

For combining just two models this approach might be adequate. However, our ultimate goal is to combine 50 or so models, with each model containing say 40

factors. Thus we could face up to 2000 factors and computing a 2000 × 2000 sample covariance matrix from limited time series data will lead to degenerate results.

In short, combining factor models has led us to exactly the same problem we faced when we tried to compute asset-by-asset covariance matrices directly. In the asset case we saw that the solution was a factor model. Hence we turn to a factor model to solve our present difficulty. We shall find a new exposure matrix $Y(t)$ and fit the model

$$f(t) = Y(t)g(t) + \phi(t)$$

Then we can form

$$\begin{aligned} G(t) &= \mathrm{cov}(g(t), g(t)) \\ \Phi(t) &= \mathrm{cov}(\phi(t), \phi(t)) \end{aligned}$$

and take as our initial estimate of $F(t)$

$$\bar{F}(t) = Y(t)G(t)Y(t)^t + \Phi(t)$$

If we suppose that the exposures in $Y(t)$ capture all sources of common covariance between $f^1(t)$ and $f^2(t)$ then $\Phi(t)$ will be a block diagonal matrix which we write as

$$\Phi(t) = \begin{pmatrix} \Phi^1(t) & 0 \\ 0 & \Phi^2(t) \end{pmatrix}$$

Let

$$H(t) = Y(t)G(t)Y(t)^t$$

which we write in block form as

$$H(t) = \begin{pmatrix} H^1(t) & H^{12}(t) \\ H^{12}(t)^t & H^2(t) \end{pmatrix}$$

Similarly we write

$$\tilde{F}(t) = \begin{pmatrix} \tilde{F}^1(t) & \tilde{F}^{12}(t) \\ \tilde{F}^{12}(t)^t & \tilde{F}^2(t) \end{pmatrix}$$

Then on the diagonal blocks

$$\tilde{F}^i(t) = H^i(t) + \Phi^i(t)$$

and on the off-diagonal block

$$\tilde{F}^{12}(t) = H^{12}(t)$$

Assuming we have found the correct factor structure, as the sample size goes to infinity one has $\tilde{F}^i(t)$ converge to $F^i(t)$. With finite samples, however, this convergence may be incomplete. We may rescale $\tilde{F}$ to bring its diagonal blocks into agreement with the blocks F^i provided by the local models. Let $M^{1/2}$ indicate the Cholesky square root of the matrix M. Introduce

$$R = \begin{pmatrix} (F^1)^{\frac{1}{2}}(\tilde{F}^1)^{-\frac{1}{2}} & 0 \\ 0 & (F^2)^{\frac{1}{2}}(\tilde{F}^2)^{-\frac{1}{2}} \end{pmatrix}$$

Then

$$\hat{F} = R\tilde{F}R^t$$

is an estimator of F such that its diagonal blocks $\hat{F}^i$ are identical with the blocks provided by the local market model. Note that as $\tilde{F}^i$ converges to F^i the rescaling matrix R converges to the identity. We take $\hat{F}$ as our final estimate of F. In particular its off-diagonal block $\hat{F}^{12}$ is

$$\begin{aligned} \hat{F}^{12} &= (F^1)^{\frac{1}{2}}(\tilde{F}^1)^{-\frac{1}{2}}\tilde{F}^{12}(\tilde{F}^{2t})^{-\frac{1}{2}}(F^2)^{\frac{1}{2}} \\ &= (F^1)^{\frac{1}{2}}(H^1 + \Phi^1)^{-\frac{1}{2}}H^{12}(H^{2t} + \Phi^{2t})^{-\frac{1}{2}}(F^2)^{\frac{1}{2}} \end{aligned}$$

To summarize the discussion, we saw that F^{12} was the information required to combine the two local market models and that its estimator $\hat{F}^{12}$ can be constructed from the given data F^1 and F^2 plus the additional data H and Φ derived from the factor model

$$f(t) = Y(t)g(t) + \phi(t)$$

Let us remark that we have described a completely general approach to combining two factor models. Nowhere did we assume that the models were in fact equity models. The technique could be used to combine two equity models. But by the same token it could also be used to combine equity models with bond or currency models. In addition, one sees easily that the technique is not restricted to combining just two models. One could combine any number of local market models $(x^i(t), F^i(t), \Delta^i(t))$ given a suitable factor model (Y, G, Φ). Thus we have in fact described a mechanism for modeling the entire universe of investable assets. The essential property of this construction is that it integrates the detailed analysis of a local market into a global analysis. In particular, if h is a portfolio which happens to lie in a local market, then the risk analysis of h conducted in the local context will be identical with the analysis of h conducted in the global framework.

Were our only goal to forecast the risk of a portfolio, then we could work entirely with $\hat{\Omega}$ and ignore the structure

$$\hat{\Omega} = X[YGY^t + \Phi]X^t + \Delta \tag{11}$$

which generates the estimate of Ω. Insight into the sources of return is also desirable, however, and so we are led to discuss the various components of (11). The introduction of some terminology permits this discussion to flow more smoothly. We have referred to (x^i, F^i, Δ^i) as the *local models* and so it is natural to speak of $f(t)$ as the *local factor returns.* Then it is reasonable to refer to (Y, G, Φ) as the *global model.* In the equation

$$f(t) = Y(t)g(t) + \phi(t)$$

we shall refer to $Y(t)$ as the *global exposures* and $g(t)$ as the *global factor returns.* This equation breaks the local factor returns down into the *global part* $Y(t)g(t)$ and the residual part $\phi(t)$ which we shall refer to as the *purely local factor returns.* We refer to Φ^i as the *covariance matrix of purely local factor returns for the* i^{th} *market.*

Suppose h is a portfolio. Then h can be written as

$$h = w_1 h_1 + w_2 h_2$$

where h_i is the subportfolio lying in the i^{th} local market and w_i is the fraction of total wealth invested in that local market. We may describe the risk exposures of h in local terms, i.e. by $w_1 h_1^t x^1$ and $w_2 X_2^t x_2$. Alternately we may describe the

risk exposures in global terms by $h^t XY$. Let $y^i = w_i h_i^t x^i$ be the local exposure vector. We may decompose the risk of the portfolio as follows

[insert diagram]

The quantity

$$h^t XYGY^t X^t h$$

which is equal to

$$y^{1t} H^1 y^1 + 2y^{1t} H^{12} y^2 + y^{2t} H^2 y^2$$

gives the total contribution to risk from global risk factors. Most local models have risk factors of different types, for instance industries and risk indices. Let $x^{i,\tau}$ be the part of x^i corresponding to exposures of type τ. Then one can decompose risk further. For instance

$$(x^{i,\tau})^t \Phi^i x^{i,\tau}$$

gives the purely local risk in market i due exclusively to type τ, whereas

$$(x^{i,\tau})^t \Phi^i (x^{i,\tau'})$$

gives the (τ, τ') cross-type purely local risk in market i. The contribution of type τ to risk in all markets will be

[insert formula]

Thus for instance, we might measure the total contribution to risk from industry exposures. Similarly we could sum all cross-type exposures.

The above risk decompositions arise organically within the framework of our model. We can easily go beyond this framework, however. Suppose we wish to decompose asset returns r in terms of a new set of asset exposures Z. In principle, one could fit the model

$$r = Ze + \psi$$

and calculate

$$\begin{aligned} E &= \mathrm{cov}(e, e) \\ \Psi &= \mathrm{cov}(\psi, \psi) \end{aligned}$$

to produce the model (Z, E, Ψ) which is required for the risk analysis. With less labor, however, we may note that the normal equations of regression theory give us

$$e = (Z^t Z)^- Z^t r$$

where $(Z^t Z)^-$ indicates the generalized inverse of $Z^t Z$. Let

$$M = (Z^t Z)^- Z^t$$

and

$$N = I - ZM$$

Note that

$$\begin{aligned} e &= Mr \\ \phi &= r - Ze \\ &= r - ZMr \\ &= Nr \end{aligned}$$

Hence

$$\begin{aligned} \mathrm{cov}(e, e) &= M \mathrm{cov}(r, r) M^t \\ &= M \Omega M^t \end{aligned}$$

and similarly

$$\mathrm{cov}(\phi, \phi) = N \Omega N^t$$

Substituting $\hat{\Omega} = XFX^t + \Delta$ for Ω we get estimates

$$
\begin{aligned}
\hat{E} &= M\hat{\Omega}M^t \\
\hat{\phi} &= N\hat{\Omega}N^t
\end{aligned}
$$

of E and Φ respectively. Thus any risk decomposition in terms of the factors Z can actually be derived from the model (X, F, Δ) without the need to explicitly estimate (Z, E, Φ). Indeed, if the factor structure X is closer to the true factor structure than Z, then the decomposition based on $(Z, \hat{E}, \hat{\phi})$ will actually give a truer picture of risk than the analysis resulting from fitting (Z, E, Φ). We refer to $(Z, \hat{E}, \hat{\phi})$ as the *emulation* of (Z, E, Φ) based on (X, F, Δ).

3.2 Local Models

Let us apply the above discussion to the formulation of a global equity analysis. Following the formulation of GEM we can decompose the return $r_{i,num}(t)$ to asset i in numeraire currency into its local excess return $r_{i,loc}(t)$ and its currency return $r_{i,cur}(t)$ as

$$
r_{i,num}(t) = r_{i,loc}(t) + r_{i,cur}(t)
$$

To achieve a global analysis we require local models for the equity asset returns and for the currency returns. Fortunately we have a large number of such models already constructed. In this section we briefly describe these models.

For the local analysis of an equity market we employ a factor structure containing two types of factors: industries and risk indices. The definition of the industry factors begins with the choice of an industrial classification system. While the precise definitions are usually chosen so as to be appropriate to the local market, the categories are sufficiently broad and distinct as to be largely similar from country to country. Once the classification system is defined, each company is assigned an industry exposure. Always the industry exposures of a company sum to not more than one. In some models exposures are zero-one variables, so a company has an exposure of one to one industry and zero to all others. In other models companies are exposed to multiple industries. In this case the industry exposures are assigned to capture the importance of that industry to the company. The assignment is usually made based on a combination of accounting data (i.e. sales in a business segment) and a style analysis of company returns against industry index returns.

The risk indices capture other aspects of a security which are useful to understanding its return pattern. Examples of such indices are measures of size, liquidity, value, yield, exposure to foreign trade, "blue chip" quality, membership in indices on which futures contracts trade, past market performance, volatility, etc. Each local model contains a set of risk indices appropriate to that market and with definitions more or less specialized to the data available in that market. The construction of the indices follows a general pattern however. First a concept is identified, e.g. size. Then specific pieces of data (known as descriptors) are identified which have a bearing on the concept, for instance market capitalization or revenues. The descriptors are then combined to form a factor which seems maximally informative about return patterns. For instance, one might take a linear combination of the descriptors where the coefficient attached to a descriptor could be interpreted as a measure of how accurately the descriptor realizes the concept. The need to handle the problem of missing data, outliers and changing accounting definitions through time tend to make the details of factor construction rather complex. Continuous factors are usually standardized against the local asset universe, a step which makes the factor exposures comparable from factor to factor. Discrete factors (e.g. a zero-one variable indicating membership in an index portfolio) are usually left unstandardized. Factor exposures are revised as new information arrives. For instance, if a company sells a division its industry exposures may differ after that point from those before that point.

Once the factor matrix X has been defined, the factor returns and specific returns are measured through the regression

$$r(t) = X(t)f(t) + \varepsilon(t)$$

Different models generally employ a weighting scheme suitable to the local market, e.g. cap weighting or GLS weights. From returns data one constructs the covariance matrices $F(t)$ and $\Delta(t)$. In some models, however, there is considerable detail to this construction. For instance, a GARCH model might be used to scale F and Δ might be based on a method similar to that described in Grinold and Drach 1992.

In investigating factor structures we have found them to differ along time and capitalization dimensions. Thus in the US, for instance, we actually estimate three separate models—a model of large cap stocks based on monthly returns, a model of large cap stocks based on daily returns and a model of small cap stocks based on monthly returns. In Europe we estimate models for each of the national markets. However, we also estimate a pan-European model. As we have noted a factor model should be fit over a homogenous group of assets. There are degrees of homogeneity and how much homogeneity a group of assets exhibits depends

in part on ones perspective. For some investors, European equities as a whole currently represent a homogenous group, while other investors follow strategies which continue to see the national markets as the primary homogenous groups.

Our global analysis will be constructed by combining local models. The wealth of local models will give us considerable flexibility in producing a global analysis. For instance, we might combine local models estimated over daily time horizons to achieve a global analysis suitable for short-term risk assessment. Alternately, we might combine local models estimated for small cap stocks to produce a global small cap analysis. In treating European assets we could choose to combine several nationally focused local models, or instead we might use our European region model. In this paper, however, we shall focus on combining local models estimated from monthly returns for large cap stocks. We focus on this particular global analysis as it is the one we can compare most directly to the GEM analysis. Currently we have local models for 22 countries which are suitable for this purpose. They are listed in Table 1, together with the time periods they cover. Together they supply coverage for some 35,000 assets, which is approximately 80% of the assets in the world equity market.

In addition to local equity models we also require a model of currency returns. Our currency model is an example of a degenerate factor model in that it has one factor per currency. In other words the exposure matrix X is the identity, so the asset returns equal the factor returns and the specific returns are zero. The subtlety of the model comes in how the factor covariance matrix is constructed. We write

$$F_{ij} = \sigma_i \sigma_j \rho_{ij}$$

where σ_i is the volatility of currency i and ρ_{ij} is the correlation between currencies i and j. The volatilities σ_i are estimated from GARCH models chosen approximately for each currency model fit over fairly high frequency data. The correlations are estimated by an exponentially weighted moving average method applied to lower frequency data.

Table 1

Country	Commencement of Factor Returns	Commencement of Covariance Matrix
Australia	Jan. 1982	Jan. 1988
Brazil	Aug. 1994	Dec. 1996
Canada	Mar. 1980	Jan. 1989
France	Aug. 1988	Sep. 1988
Germany	Jul. 1984	Jul. 1985
Greece	Jan. 1991	Jun. 1996
Hong Kong	Jan. 1988	Dec. 1996
Indonesia	Feb. 1993	Jan. 1997
Japan	Apr. 1978	May 1983
Korea	Feb. 1986	Apr. 1986
Malaysia	Sep. 1991	Jan. 1996
Mexico	Mar. 1992	May 1992
Netherlands	May 1985	Jan. 1994
New Zealand	Jul. 1988	Jan. 1990
Singapore	Jan. 1991	Jan. 1997
South Africa	Jan. 1993	Jan. 1996
Sweden	Feb. 1984	Jan. 1989
Switzerland	Jul. 1984	Jul. 1989
Taiwan	Feb. 1984	Apr. 1994
Thailand	Oct. 1988	Apr. 1994
United Kingdom	Jan. 1981	Jan. 1981
United States	Jan. 1973	Jan. 1973

3.3 The Equity Covariance Structure

3.3.1 Methodology

Given the local models, our task now is to construct the global model which will unify them. Hence the first order of business is to determine the global factor structure Y. We focus first on the factors relating local equity markets. The research described in section two has already identified a number of possible factors

1. country, region and world factors
2. industry and sector factors
3. regional industry and regional sector factors
4. regional and global risk indices

In addition, we shall define various macro-economic factors. Give the abundant supply of factor concepts our task is to screen these possibilities to find the significant sources of risk and then to combine them into a parsimonious model.

As a preliminary screening of factors we proceed as follows. Let $f(t)$ be the vector of equity factor returns and let $y(t)$ be a vector of factor exposures. For each time period t we perform the equal weighted regression

$$f(t) = \alpha(t) + \beta(t)y(t) + \varepsilon(t)$$

Let $R^2(t)$ be the R^2 from one such regression and let $T(t)$ be the T-statistic on $\beta(t)$. We repeat this regression for $t = 1, 2, ...n$ thus building up time series of $\beta(t), R^2(t)$, and $T(t)$. We can characterize the factor $y(t)$ in several ways. First,

$$\bar{R}^2 = \frac{1}{n}\sum_{t=1}^{n} R^2(t)$$

gives the average fraction of variation in $f(t)$ explained by the factor. Second,

$$F = \frac{1}{n}\text{card}\{t : |T(t)| \geq 2\}$$

tells us how often the factor is a significant contributor to variation in return. Third, the time series mean μ and variance σ^2 of $\beta(t)$ describe the trend and

variability of this factor return. Note that these measures are sensitive to the scaling of $y(t)$. Fourth, let the cumulative factor return be

$$B(t) = \sum_{s=1} \beta(s)$$

and take

$$M = \max\{B(t)\} - \min\{B(t)\}$$

as a measure of how marked trends in $B(t)$ have been over subperiods. Fifth, the serial correlation

$$C = \operatorname{cor}(\beta(t-1), \beta(t))$$

measures the tendency of $\beta(t)$ to trend. Sixth the serial correlation

$$H = \operatorname{cor}(\beta(t-1)^2, \beta(t)^2)$$

in $\beta(t)^2$ is sensitive to the conditional heteroscedasticity in $\beta(t)$. Broadly, the measures R^2, F, σ^2 and H tell us how important the factor is for risk, while the measures $\mu, M,$ and C tell us the importance of the factor for alpha. A factor which is important for alpha, even if non-descript from a risk viewpoint, is likely to attract active strategies and so from an application standpoint it may be important to characterize risk along such a dimension.

A seven dimensional characterization of a factor is somewhat awkward to work with. As a convenience we create a score variable to summarize factors which appear important by one or more characterizations. Specifically for each characterization i we define S_i to be 0, 1 or 2 if the factor appears to be of little, some or significant importance when judged by characteristic i. Then the total score for the factor is

$$S = \sum_i S_i$$

As an example the T statistic would be expected to have a value of 2 or more 5% of the time by chance. Thus we would set the score variable for F, the fraction of the time the T statistic exceeded 5%, as

$$S_F = \begin{cases} 0 & F \in [0, 0.05) \\ 1 & F \in [0.05, 0.1) \\ 2 & F \geq 0.1 \end{cases}$$

For the other statistics we get the score as 0, 1 or 2 depending on whether the statistic lies in the lower middle or upper third of the range of observed values across the set of factors. It will be seen that there is no particular significance to the score statistic. It simply serves as a convenience in data summarization.

3.3.2 Factor Screening

We take ??? as our study period. We will examine factors in related groups, e.g. all countries together, then all industries together, etc. The first group we examine is countries. The most natural exposure of a factor to a country factor would be the factor beta, i.e. the beta from a regression of the factor returns on the market return. As a first approximation to the factor beta we set

$$y_i^j(t) = \begin{cases} 1 & \text{if } i \text{ is an industry in country } j \\ 0 & \text{otherwise} \end{cases}$$

Note that the factor return to y^j is just the equal weighted mean industry return in country j. We test each y^j separately and present the results in Table 2. It will be noted that judged by the T statistic every factor is important. This is an expected result, consistent with the findings of previous researchers that countries represent an important dimension of risk.

Next we examine industries. For this purpose we use the industry classification shown in Table 3. We assign each local industry to exactly one of these industries. We define the exposure vectors y^j by

$$y_i^j = \begin{cases} 1 & i \text{ is an industry which is assignd to the } j^{th} \text{ industry in Table 3} \\ 0 & \text{otherwise} \end{cases}$$

Table 3: Industries

Aerospace	
Airline	
Autos	
Bank	
Business services	
Chemicals	
Commerce	
Computers	
Construction	
Dstg	Distribution
Electronics	
Energy	
Entertainment	
Equipment	
Financial	
Food	
Health	
Industrials	
Information technology	
Insurance	
Media	
Metals	
Mining	
Miscellaneous	
Oil	
Precious metals	
Property	
Raw materials	
Retail	
Services	
Technology	
Telephone	
Transportation	
Utilities	

The result of testing the y^j is shown in Table 4. It will be noted that very few industries have a significant T statistic. To refine the test we modify our procedure. We know country factors account for a significant amount of noise in the data. To eliminate this source of variation we replace the local industry factor returns by their active values (i.e.we subtract off the mean industry factor return within the local market, thus removing the country effect). The results of testing the industry effects against the active local factors is shown in Table 5. Now the industry effects emerge from the noise. Clearly not every industry is an important factor. We will retain for further study only those factors, marked by an asterisk in Table 5, which seem of possible importance.

Next we test for regional effects. The regions we shall consider are defined in Table 6. We set y_i^j to one if factor i is an industry factor in a country belonging in region j. We test the regions against the full factor returns, with the results as shown in Table 7. We find evidence of a world, Europe, Latin America and Asia ex Japan factors. By contrast, evidence for an emerging market factor is weak. We note that this result is consistent with a study of van Royen ["Measuring Financial Contagion: Evidence from the Portfolio Flows of International Investors", Proceedings of the Q Group Spring 2000 Seminar] which examined portfolio flows of international investors to see if there was a correlation in flows. This study defined its regions as Latin America, Asia ex Japan, Europe Periphery (Greece, Hungary, Poland, Turkey, UK), Commodity Linked (Australia, Canada, New Zealand, Norway, Sweden) and Safe Havens (core Europe, Japan). Linkages across these regions were found to be weak.

Table 6: Regions

Name	Constituents
ANZAK	Australia, New Zealand
China	Hong Kong, Taiwan
Developed 1	Japan, US, Canada, Europe 1
Developed 2	Japan, US, Canada, Europe 2
Developed 3	Japan, US, Canada, Europe 3
Emerging	Brazil, Mexico, Tiger
Europe 1	Europe 3, Switzerland, Greece, Sweden
Europe 2	Europe 3, Switzerland, Sweden
Europe 3	France, Germany, UK
North America	US, Canada
Pacific	Japan, ANZAK, Tiger
South America	Mexico, Brazil
SNML	Singapore, Malaysia
Tiger	Taiwan, Hong Kong, Indonesia, Malaysia, Singapore, Thailand, Korea

For the defined regions we consider regional industries, i.e. we take $y_i^{j,k}$ to be one if i is a factor from a market in region j and i belongs to industry k. Table 8 shows the evaluation of these factors for the active factor returns (i.e. with country effect removed.)

Similar to the agglomeration of countries into regions we may consider industries mapped into sectors. We employed the industry-to-sector mapping devised by Hui and given in Table 9. The results for global and regional sectors, evaluated on local returns net of country effect, are given in Tables 10 and 11.

Table 9

Sector	Constituents
Basic Materials	precious metals, metals, chemicals, mining, raw materials
Industrial 1	equipment, construction, industrials
Transportation	airline, transportation
Technology	telephone, computer, technology, aerospace, electronics, information technology, business services
Financial	bank, insurance, property, financial services
Energy	energy, oil
Utilities	utilities
Consumer Staples 1	food
Health	health
Consumer Cyclical	arts, retail, media, entertainment
Commercial Services	commercial services
Industrial 2	industrial, transportation
Consumer Staples 2	consumer staple 1, health
Consumer Cyclical 2	consumer cyclical 1, commercial services
Industrial 3	basic materials, industrial 2, technology
Consumer	consumer staple 2, consumer cyclical 2

Next we consider macroeconomic variables. Here the question is whether countries whose economies are similar from a macroeconomic viewpoint exhibit similar patterns of equity returns. Thus these variables may be considered as an alternate characterization of country and regional effects. For each country we define 10 macroeconomic descriptors:

Concept	Definition
per capita income	GDP in $/population
importance of equity market	market cap/GDP
labor stress	unemployment rate
monetary stress	consumer price inflation rate
misery	labor stress & monetary stress
financial stress	risk-free rate
trade balance 1	current account/GDP
trade balance 2	current account/reserves
trade balance 3	current account/foreign exchange
importance of exports	exports/GDP
importance of trade	(exports & imports)/GDP

Then to each industry factor we assign the exposure of its country descriptor. We again assign risk indices a zero factor exposure descriptor. When these factors are evaluated on the full local factor returns they initially appear quite significant (Table 12). However, when they are re-evaluated on the local factor returns net of country effects most of this explanatory power disappears (Table 13). Thus, we concluded that the macroeconomic data in general adds little to the country effects. We made some further analysis whose details it would be tedious to persue and concluded that we could devise one macroeconomic factor of some interest above and beyond country effects. We term this factor EQUILIBRIUM. It is the sum of two descriptors. The first descriptor is the country's inflation rate standardized cross-sectionally. The second descriptor is the ratio of the country's trade balance to its reserve bank holdings of foreign exchange, also standardized cross-sectionally. [3] Thus this factor is sensitive to either internal or external monetary disequilibrium.

Next we consider the risk indices. There are several classes of risk indices which are prevalent in the local models (Table 14.) We assign factor i the exposure $y_i^j = 1$ if i belongs to the j^{th} risk index type and otherwise zero. Testing these factors against the active factors (i.e. net of country effect) we get the results shown in Table 15. We also tested various regional factors, i.e. $y_i^{j,k} = 1$ if i is a factor of type j for a country in region k and otherwise $y_i^{j,k} = 0$. These results are shown in Table 16.

[3] Here trade balance is imports minus exports.

Table 14: Risk indices

Book-to-price
Capital intensity
Currency sensitivity
Earnings variability
Export orientation
Foreign sensitivity
Growth
High value
Index constituent
Interest rate sensitivity
Labor intensity
Non-estimation universe
Profitability
Size STIL
Value
Variability in markets
Volatility
Yield

Surveying these results we are in a position to make an initial judgment about factor structure. We decided to retain all the countries, the marked industries, the marked risk indices and EQUILIBRIUM for second stage analysis. We set aside for the moment the regional, regional industry, sector, regional sector and regional risk index factors at this point as they appear to add little value over the already selected factors.

3.3.3 Model Selection

So far we have tested each factor in isolation. The next step is to test the importance of factors jointly. We use the following algorithm to make this measurement.

1. Let $T = \{f_1, f_2, ... f_n\}$be the set of factors to be tested. Let S be the set of factors which have already been tested. Initially $S = \{\ \}$, the empty set. Let m be a counter variable with $m = 1$ initially.
2. For each f_i in T let S_i be the result of adding f_i to S. For each t fit the model which regresses the local factor returns against the variables in S_i. Let the R^2 be denoted $R_i^2(t)$.
3. Form the quantities

$$\bar{R}_i^2 = \frac{1}{n}\sum_{t=1}^{n} R_i^2(t)$$

which measure the mean R^2 of the model S_i. Select that i for which $\bar{R}_i^2$ is greatest. Let Q_m be this $\bar{R}_i^2$. Remove f_i from T and add f_i to S.

4. If T is empty, i.e. all factors have been tested, then stop. Otherwise let $m = m + 1$ and go to step 2.

In this way we add factors to the model in such a way that the incremental fit improves maximally at each step. Table 17 lists the factors in their order of addition, together with Q_m and $\Delta Q_m = Q_m - Q_{m-1}$. We take ΔQ_m as a measure of the marginal improvement in fit due to the m^{th} factor added. This is a conservative measure of the marginal contribution of a given factor in the sense that all other factors have been given an opportunity to enter the model ahead of the tested factor and so reduce its measured marginal contribution to fit. Figure 1 shows a graph of ΔQ_m against m. Rather arbitrarily we cut-off the model building once ΔQ_m drops to 0.1. This results in a 48 factor model as shown in Table 17.

As a further refinement of the model we consider the influence of different weighting schemes. Table 18 shows the impact on mean R^2 and root mean square error of using equal weights, cap weights, log cap weights, GNP weights and GLS weights. Based on this finding we selected GLS weights.

3.3.4 Construction of the Equity Covariance Matrix

Our next task is to generate a covariance matrix G of global factor returns. Here we face the difficulty that our local models begin at different dates. In consequence the country factor returns which we are able to estimate have different starting dates. It is desirable that all the time series entering the covariance matrix estimation have the same length. Accordingly we need to extend some of the country factor return series backwards in time. Our solution is to proxy the missing country factor return data with data derived from a local market return index. Once all the time-series have been completed to the same length we form the exponentially weighted expanding window sample covariance matrix. In other words the matrix estimated at time t uses all information from periods prior to t with data n periods in the past being weighted by α^n for α a suitable constant. In fact we pick α to have a 90 month half-life, i.e. $\alpha^{90} = 0.5$.

With such a slowly decreasing half-life, our covariance matrix is not particularly responsive to current conditions. To cure this difficulty we employ a GARCH scaling technique. We take r_t to be the returns to a cap weighted world market index measured weekly and we fit the model

$$\begin{aligned} r_t &= \alpha + \varepsilon_t \\ \varepsilon_t &\sim N(0, h_t^2) \\ h_t^2 &= \omega + \alpha h_{t-1}^2 + \beta \varepsilon_{t-1}^2 + \gamma \varepsilon_{t-1} \end{aligned}$$

We select a weekly measurement framework to minimize problems of non-synchronous trading which are unavoidable in world indices while maximizing the number of data points. The coefficients of the model we estimate are given in Table 19. The time-series graph of h_t is shown in Figure 2. This model gives us an estimate of volatility over a weekly horizon. We aggregate it to a monthly horizon thus arriving at the model

$$h_{month}^2(t) =$$

Let G_t be the global factor covariance matrix formed previously. We form its principal components decomposition, i.e. we write

$$G_t = O_t D_t O_t^t$$

for O_t an orthogonal matrix and a diagonal matrix D_t whose diagonal components $d_i(t)$ satisfy

$$d_i(t) \geq d_{i+1}(t)$$

We shall make two modifications to D_t. First, we identify the first principal component with the world factor and replace $d_1(t)$ by $h_{month}^2(t)$ calculated from the GARCH model. Next we find the number p such that the p smallest diagonal components of D_t are not statistically distinguishable. We replace these components by their mean. The motivation for this replacement is that these components correspond to different dimensions along which risk can vary. Since, however, we do not believe we can estimate meaningful differences in risk, we prefer to set the estimate to a mean value so that portfolio construction is not driven by spurious differences. Let $\tilde{D}_t$ denote the modified diagonal matrix. Let

$$\tilde{G}_t = O_t \tilde{D}_t \, O_t^t$$

then $\bar{C} = YGY^t + \Phi$ is our estimate of the equity part of the global covariance matrix.

3.4 The Currency-Equity Covariance Structure

Having disposed of the equity part of the global model we turn to the currency part. There are two sets of covariances. We must estimate currency-to-currency covariances and currency-to-equity covariances. The currency-to-currency part is immediately disposed of by our currency model. Hence it is enough to consider covariances between currency returns and local equity factor returns. As usual the dimensions preclude a simple sample covariance matrix. Looking cross-sectionally we find that in most countries the important currency covariance is with either the market factor or the size factor. For each market we focus on the currency returns in terms of its local currency. Currency returns have two interpretations. First, as economic variables they capture important influences on the local economy. From this perspective one would expect important correlations with major trading partners or with important currencies in international trade (e.g. the dollar.) Second, currencies are sensitive indicators of political stability. When the stability of the local market is of concern then the exchange rate with a major currency such as the dollar will measure the perceived risk. Alternately the political risk may be in a third country, in which case that country's exchange rate, either to local or to a major currency may be the relevant measure of risk. On examining the record we find a number of local markets exhibiting a correlation with the Russian ruble exchange rate. This correlation in general stems from the Russian default of August 1998 and its aftermath. This finding calls for an application of analyst judgment. On the one hand, one could treat this event as a one time event which is unlikely to reoccur and thus treat this correlation as uninformative as to future correlations. On the other hand, one could conclude that Russia's ability to disturb the world capital market stems from its combination of military power and internal instability. These characteristics are intact and we may anticipate that Russian developments will episodically impact the capital markets in the future. Surveying the pattern of correlation with the ruble (Table 20) we generally find that Russia's European neighbors and the emerging market countries display sensitivity to this factor, whereas countries such as Canada and the US do not. This pattern is consistent with the concept that the ruble correlation is capturing an element of political risk. Accordingly, we have decided to treat this correlation as an enduring risk factor. To improve the interpretability of this factor, we filter out the low level variation in the currency return. Specifically let $r(t)$ be the ruble-dollar exchange return. Let $\sigma(t)$ be the sample standard deviation of $r(t)$. We set

$$c_1(t) = \begin{cases} 1 & \text{if } |r(t)| > 2\sigma(t) \\ 0 & \text{otherwise} \end{cases}$$

Similarly we take $c_2(t)$ to be the dollar-to-local return, $c_3(t)$ to be the Euro-to-local return and $c_4(t)$ to be the Pound-to-local return[4]. Take $f(t)$ to be the local equity return (either the market index or size.) We find the best one factor time-series model of the form

$$f(t) = \alpha + \beta c_i(t) + \varepsilon(t)$$

We then seek the best two factor model based on the quantities $c_i(t)$. In all cases we find that a second factor exhibits negligible improvement in fit. In this way we fit one factor models linking equity and currency markets (Table 21.) Let

$$f(t) = \alpha + \beta c(t) + \varepsilon(t)$$

be the fitted model for some market. If $\tilde{f}(t)$ is a second equity factor in that market and $\tilde{c}(t)$ is a second currency factor we assume that the covariance between $\tilde{f}(t)$ and $\tilde{c}(t)$ derives from the linkage through $f(t)$ and $c(t)$. Then

$$\begin{aligned} \tilde{f}(t) &= \alpha_1 + \beta_1 f(t) + \varepsilon_1(t) \\ \tilde{c}(t) &= \alpha_2 + \beta_2 c(t) + \varepsilon_2(t) \end{aligned}$$

and thus

$$\begin{aligned} \mathrm{cov}(\tilde{f}, \tilde{c}) &= \beta_1\beta_2\mathrm{cov}(f, c) \\ &= \beta_1\beta_2\beta\mathrm{var}(c) \end{aligned}$$

Here only the quantity β is actually estimated. The quantities β_1, β_2 and $\mathrm{var}(c)$ are calculated from the equity and currency covariance matrices respectively. In this way we may compute the covariance block between the currencies and that local equity market.

[4]The Euro is proxied by the Deutschmark prior to date.

3.5 The Total Covariance Structure

Let $i = 0$ indicate currencies and for $i > 0$ let i denote the i^{th} local equity market. We have described the construction of the covariance block $C^{i,j}$ giving the covariances between the factors of the i^{th} and j^{th} markets. We assemble these blocks into a large covariance matrix $\tilde{F}$ where

$$\tilde{F} = \begin{pmatrix} C^{0,0} & C^{0,1} & C^{0,2} & \cdots & C^{0,n} \\ C^{1,0t} & C^{1,1} & C^{1,2} & \cdots & C^{1,n} \\ \vdots & \vdots & \vdots & & \vdots \end{pmatrix}$$

This is our preliminary estimate of the common factor covariance matrix. We rescale it as described in section 3.1 to bring the diagonal blocks to their target values, thus achieving our final estimate of common factor covariance matrix $\hat{F}$. This completes our construction of the global equity model. It is convenient to have a name by which to refer to this model. We indulge in some word play and name it Solitaire. [5]

4 Properties and Applications of the Model

Having formulated Solitaire, it is of interest to probe the model's properties and to apply this new tool to some questions in global equity analysis.

4.1 Properties of the Global Model

We focus first on properties of the global model. Figure 3 shows the monthly R^2 of the model, together with a 6 month moving average. While model fit varies from month to month, average fit is stable. Summary statistics on fit are given in table 21. Note that the average level of fit is —, as compared with a typical value of 30% for single country equity factor models. The higher fit in the global model can be understood as the model is fit at the factor return level rather than at the asset level. Figure 4 plots monthly fit against the absolute return to the world factor. Just as the fit of local models improves in months when there is a strong local market effect, so the fit of the global model improves in months when there is a strong world market effect. In fact, the regression

$$R^2(t) = \alpha + \beta|w(t)| + \epsilon(t)$$

[5] Solitaire = a large multi-faceted gem, usually a diamond = name of the tarot card reader in Ian Fleming's "Live or Let Die" = what is left of a bridge game after the partners leave.

on the absolute world return $w(t)$ has the properties shown in table 22. In particular the magnitude of the world return explains —% of the variation in the fit of the global model.

In the construction of factor models a point of concern is that multicollinearity in the factor exposures may disturb the factor returns, so leading to an exaggerated estimate of variance. We control for this possibility by computing the variance inflation factors (VIFs) of the model. There is one VIF for each of the factors entering the model. Figure 5 plots the highest value of these VIFs for each month. It can be seen that the VIFs are small in magnitude and of stable behavior. Hence we conclude that multicollinearity is not a concern.

A key assumption of our model is that the global factor structure Y accounts for all significant covariances between local models. An implication is that the sample covariance matrix

$$\Phi_{ij}(t) = \mathrm{cov}(\{\phi_i(u)\}_{u=1}^{t},\ \{\phi_j(v)\}_{v=1}^{t})$$

should be block diagonal. We test this hypothesis by computing the mean of the cells of the off-diagonal blocks in Φ and also the mean magnitude of these off diagonal cells. For the whole period we find a mean of —- and a mean magnitude of —-. Figure 6 shows these quantities calculated from a 60 month moving window. The low level of correlation outside the diagonal blocks and the stability of this behavior supports the hypothesis that Y captures the important sources of cross country correlation.

To summarize the discussion to this point, the global model fits well, appears free of degenerate behavior and conforms to the posited model structure.

4.2 Properties of the Global Factors

We turn from a consideration of the model as a whole to a consideration of the properties of its factors. Table 22 summarizes the important time-series properties of the model's factors. Graphs of all the factor returns are given in the appendix. Here we call attention to three of the more interesting ones: ENERGY, SIZE and EQUILIBRIUM (figures 7-9). Comparable to work done at the asset level, we may also calculate the marginal contribution to fit of the different types of factors (table 23). We find the order of importance of the factors to be world, country, industry and risk index when measured at the group level. Allowing for the number of factors in each group, however, we do not find the differences between countries, industries and risk indices to be so marked, however. Some insight into the factor covariance matrix can be gained from its largest variances and correlations (table 24.)

4.3 Properties of the Purely Local Factors

A new analytical concept introduced with the model is that of the purely local factor return. We may investigate the structures which it reveals in several ways, First, for each country i the ratio det Φ^i/ det F^i gives a measure of how segmented that market is within the global market. Table 25 tabulates these ratios. We saw how in the investigations of Chaumeton and Coldiron and of Chandrashekaran, Hui and Rudd an effort was made to distinguish between domestic and global companies and thus to distinguish between markets based on the fraction of market capitalization accounted for by each type of asset. Here we have a similar division into local and global, in this instance achieved at the factor level. Table 25 provides a comparison of our characterization with the other two.

We may also study the pattern of purely local returns from a factor viewpoint. For each factor i we may take the ratio var(ϕ_i)/ var(f_i) as a measure of how local that factor is. Collecting the means of these ratios over factors of a given type, e.g. TECHNOLOGY or SIZE, we get a measure of how much local variation in behavior there is within the scope of a globally defined concept. Here the results are presented in table 27.

There is an important economic interpretation which can be put on the purely local factor return. Within a market these factor returns represent common risk factors which cannot be eliminated through diversification. For investors who operate globally, however, the purely local factor risk can be reduced by diversification across markets. The implications are two. First, markets where detΦ^i/ detF^i are high are the markets in which the benefits of international diversification are greatest. Second, the trend in this ratio through time shows how the benefits of international diversification have varied. In table 28 we present the results of the regression

$$\det \Phi^i(t) \ / \ F^i(t) = \alpha + \beta t + \epsilon_i(t)$$

and in figure ~~?~~ 10 we graph the ratio through time for selected countries.

4.4 Portfolio Analysis

4.4.1 Market Portfolios

Next we turn to a consideration of portfolios. The simplest portfolio to consider is the market portfolio h_m^i of country i. The analysis of this portfolio in Solitaire is identical with its analysis in the single country model for that country. By contrast, the analysis in GEM is different. Table 29 compares the Solitaire estimate of risk with the GEM estimate.

The next level of analysis considers the covariance between h^i_m and h^j_m for two countries i and j. In Solitaire this measure probes the off diagonal block C^{ij}. Table 30 presents the Solitaire and GEM analysis. It will be seen that Solitaire typically estimates a lower correlation, which implies a greater benefit from international diversification. Here the essential difference in the two analyses appears to be Solitaire's recognition that much cross-market correlation derives from the world factor which displays pronounced conditional heteroscedasticity, whereas GEM assumes a homoscedastic world. If we drop from the empirical data the months when the world return was greatest (10/87, 8/98) and calculate the correlations of the empirical data series, we recover values which are closer to the Solitaire estimates than to the GEM estimates (Table 30.)

4.4.2 A Case Study

To gain further insight we consider a simple case study. We take our investable universe to consist of just four stocks: Toyota, Mitsubishi Bank, General Motors and Citicorp. We will take the benchmark portfolio to be the equally weighted portfolio holding these four assets. We take the managed portfolio to consist equally of Mitsubishi Bank and General Motors. Thus the active portfolio (i.e. the difference between the managed portfolio and the benchmark) is

Toyota	-25%
Mitsubishi Bank	25%
General Motors	25%
Citicorp	-25%

It will be noted that the active portfolio has no net exposure to Yen/Dollar, Japan/US or Autos/Banks. We analyze the risks of the active portfolio in both GEM and Solitaire. In GEM the risk exposures are

[insert table]

It will be noted that GEM perceives a country risk, despite there being no active distribution in portfolio wealth between countries, as GEM calculates the country exposures based on historical beta and the betas of these assets are not all 1.0. In fact, the computation of the GEM country exposure is detailed in table 31. In figure 8 we present the GEM risk decomposition. Most of the risk is specific. Such common factor risk as there is derives largely from the country exposures. Next we turn to the Solitaire analysis. Like GEM, Solitaire sees no currency exposure in the active portfolio. In other details, however, Solitaire paints a significantly different picture. The major risk exposures revealed by Solitaire are given in table 32. Two exposures in particular are revealed which

the GEM analysis missed. First, the Mitsubishi Bank is more domestically focused than Toyota, so the active portfolio contains a significant tilt towards the Japanese domestic economy as revealed by the foreign sensitivity risk index. Second, General Motors has significant exposure to the financial service industry through its credit corporation subsidiary, a fact which is captured by the multi-industry exposures supported by the USE3 model underlying Solitaire. The Solitaire risk decomposition is shown in figure 9. Superficially the Solitaire analyses seems similar to the GEM analysis in that the estimated total active risk is nearly the same in the two analyses (—— vs ——). Probing more deeply, however, we find striking differences in the risk decomposition. The common factor risk estimated by Solitaire is ——, nearly three times the GEM estimate of ——. Most of this common factor risk derives from the Japan subportfolio, and may be attributed nearly equally to industry tilts and to risk index tilts. The low value of cross-covariance between the US and Japan subportfolios (——) indicates that relatively little hedging occurs between those portfolios. GEM assumed that the industry tilts within the US and Japan subportfolios canceled one another. Solitaire is sceptical of this cancelation. In fact the low risk of the US subportfolio seems to derive from General Motor's exposure to both autos and financial services.

4.4.3 A Second Case Study

Let us consider the implication of this study. As we know, the specific risk of a portfolio (measured in variance term) will scale like $1/N$ for N the number of assets in the portfolio. The common factor risk, however, will generally not decrease with rising N. As we have seen, GEM and Solitaire differ in their perception of how much common factor risk a portfolio contains. By allowing N to increase, we should be able to dramatize the difference in the two analyses. For this purpose, we take our investable universe to consist of banks and chemical companies in the US and Japan respectively. We switch from autos to chemicals as the chemical industry is less consolidated in the two countries. We take the benchmark portfolio to be allocated as

25%	Japan Banks
25%	Japan Chemicals
25%	US Banks
25%	US Chemicals

and to be cap weighted within each subdivision. The managed portfolio holds the Japan Banks and US chemicals, resulting in the active portfolio shown in table 33. The GEM and Solitaire risk decompositions are given in figures 10 and 11. As expected, the analysis is now dominated by common factor risk and

Solitaire estimates a total risk approximately twice that of GEM. As noted in section 3.2 it is possible for Solitaire to emulate the GEM risk decomposition. The comparison of the actual GEM decomposition and the Solitaire emulation is given in table 33. Note that the Solitaire emulation puts the various risk components in the same proportions to one another as the GEM decomposition, but that Solitaire scales the components up to reflect the higher level of total active risk estimated by Solitaire.

4.5 Accuracy of Risk Forecasts

As we have seen, there are important differences between the GEM and Solitaire risk analyses. Thus, it is natural to ask which analysis is more accurate. We may evaluate the accuracy of a risk forecast as follows. For some portfolio, let $r(t)$ denote its return over the period starting at time t. Let $\sigma(t)$ be the forecast standard deviation of return over the period starting at t, with the forecast being made at the start of the period. Form the standardized return series

$$z(t) = r(t)/\sigma(t)$$

Ideally we would like the $z(t)$ to be independent identically distributed random variables with a standard deviation of 1. We can test this hypothesis by fitting the GARCH model

$$\begin{aligned} z(t) &= \mu + \hat{h}(t)\, \epsilon(t) \\ \epsilon(t) &\overset{\sim}{iid} \cancel{T_n}\ N \\ h(t)^2 &= w + \alpha\, h(t-1)^2 + \beta\, \overset{\varepsilon}{\cancel{h}}(t-1)^2 + \gamma\, \overset{z(t-1)}{\cancel{h(t-1)\,\epsilon(t)}} \quad z(t-1) \end{aligned}$$

Here T_n is the T distribution with n degrees of freedom. We employ the T distribution as we allow the possibility that the tails of the $z(t)$ distribution are heavier than the normal. If our hypothesis is true (i.e. $z(t)$ is i.i.d with standard deviation one) then we should find that the hypothesis $\alpha = \beta = \gamma = 0$ cannot be rejected and that the unconditional variance of the fitted series, namely

$$\frac{w + \mu\gamma}{1 - \alpha - \beta} \qquad \mu = 1$$

is one. Table 36 exhibits the results of testing both GEM and Solitaire over a range of portfolios.

5 Assessments and Future Directions

At this point we have gathered sufficient data to permit a preliminary evaluation of the integrative modeling approach. Fundamentally, we have shown how exposure based factor modeling may be applied to the problem of combining local factor models. This step completes the program initiated by Hui of constructing a global risk analysis from local models. The result we call Solitaire. The expected advantage of the new approach is the preserving the scope of the old global analysis (GEM) while achieving the depth of analysis inherent in the underlying local models. We have seen in detail in the four asset case study how Solitaire permits us to probe more deeply into the risks of an equity portfolio. From these deeper insights we expect superior risk forecasts. Our tests of risk forecasting accuracy indicate that meaningful differences in risk are forecast by GEM and Solitaire and that the Solitaire forecast is generally superior. Thus our program has achieved some success in the domain of risk analysis.

Beyond risk analysis, Solitaire also provides us with a new vantage point from which to survey some of the classic questions in global equities. Our results broadly confirm prior research in that risk is seen as derived from country, industry and risk index risk exposures. The introduction of EQUILIBRIUM and the heteroskedastic world market factor somewhat expand the treatment of risk factors, however. Perhaps more significantly we introduce the concept of a purely local common factor. This concept sheds new light on the benefits of international diversification. It also provides a new angle on the question of domestic versus global in the context of national markets, a question which has previously been grappled with by other authors at the asset level.

Despite these advances, some limitations of Solitaire should be noted. Currently the model covers only 80% of global equities by name. Historical coverage is somewhat ragged, with important markets (particularly in Europe) not having the length of coverage which we might desire. These circumstances somewhat limit the tests of Solitaire which we have been able to perform. More extensive out of sample testing in particular is to be desired. In general testing with optimized portfolios provides the most stringent measure of risk model performance. With Solitaire currently only partially integrated into our optimization framework, the degree of testing we have been able to do here is necessarily somewhat limited.

Removing these limitations provides the immediate focus for our ongoing research efforts. Given progress already made, we expect to extend Solitaire's coverage to the entire equity universe in short order, and with backfilling of history, we anticipate having historical coverage comparable to GEM. These two

achievements will then make possible a comprehensive program of model testing which should better define the accuracy of the Solitaire forecast and possibly lead to refinements in the model.

Looking further out, we can see that Solitaire carries implications for our total equity research program. If single country models are to be joined together to form a global analysis, then it is desirable that the features of the local models should be as regular as possible. For instance, it is desirable that local industry definitions differ from the global industry structure only when well justified by the local situation. Cleaning up this "definitional noise" will permit more accurate evaluation of the true strength of the purely local factor returns. Conversely, for the modeling of the smaller local markets, an implication of Solitaire is that global factor data may be applied to establish appropriate priors. Finally we have noted that Solitaire may be developed in small cap, short-term and regional versions. Realizing this program will permit cross-comparisons between the different versions, which - one may anticipate - will shed considerable light on the functioning of the global equity market. Looking within a single equity market, we can see now the technique of combining factor models could be applied at the sector level to achieve an even deeper analysis of local markets. For instance, one might model US equities by integrating separate factor models for Industrials, Financials and Service Companies. In the global context a factor model for multi-national companies might be developed and integrated with the geographically based local equity models to develop an enhanced global analysis. Looking beyond just equites, we see that Solitaire points the way to deepening our analysis of cross-asset class risk. Clearly then Solitaire provides a framework which will guide much of our research in coming years.

Glenn Pat[redacted]up
3475 Ediso[redacted]Suite L
Menlo Park, CA 94025

Tel 650 474 8400
Fax 650 474 8401
glenn-law.com

G|P|G

January 29, 2001

Leslie Santos
BARRA, Inc.
2100 Milvia Street
Berkeley, CA 94704

Re: Patent Application Entitled AN INTEGRATIVE METHOD FOR MODELING THE WORLD EQUITY MARKET
Our File: BARR0005

Dear Leslie:

Enclosed for your review is a draft of the above-identified patent application together with informal versions of our proposed drawings. Please have the inventors review the application to ensure that:

(1) it is accurate and complete in its descriptions;
(2) it sets forth sufficient detail to enable those skilled in the art to make and use the invention;
(3) the best known mode of practicing the invention, (i.e., the preferred way of making and using the invention is disclosed).

To the extent possible, please have them make any changes to the specification on this draft as they would like them to appear in the finalized application. Then when they have finished reviewing the application, please return the marked-up version of the draft. After we have received their comments and made any appropriate revisions, we will prepare the final application, which will be forwarded to you for your review.

If the application is in good order as enclosed or with the inventor's minor additions, please contact me with the following information so that I may prepare the signature forms.

- All inventor's full names, residence addresses, and country of citizenship

DOCKETED DRWS
By:
Date:

- Whether you want an assignment prepared; I'll need the full name, address, and state of incorporation for the company
- Does the company have less than 500 employees, if so, who (name and title) is authorized to sign on behalf of the company.

As you are aware, we have a duty to disclose the most pertinent prior art of which you are aware to the Patent Office. If you can think of any pertinent references or patents, or simply any similar existing technology, please let us know, so that we can include a discussion of them with an Information Disclosure Statement. Additionally, please remember that the duty to disclose pertinent prior art continues until the patent actually issues. Therefore, if you become aware of other prior art designs in the future, please let us know.

Very truly yours,

Jessica Pallach

Jessica Pallach
Patent Administrator

/jlp

Enclosures

Glenn Patent Group
3475 Edis[illegible] Suite L
Menlo Par[illegible]94025

Tel 650 474 8400
Fax 650 474 8401
glenn-law.com

GPG

April 24, 2001

Leslie Santos
BARRA, Inc.
2100 Milvia Street
Berkeley, CA 94704

Re: Patent Application Entitled AN INTEGRATIVE METHOD FOR MODELING MULTIPLE ASSET CLASSES
Our File: BARR0005

Dear Leslie:

Enclosed for your review is a draft of the above-identified patent application together with informal versions of our proposed drawings. Please have the inventors review the application to ensure that:

(1) it is accurate and complete in its descriptions;
(2) it sets forth sufficient detail to enable those skilled in the art to make and use the invention;
(3) the best known mode of practicing the invention, (i.e., the preferred way of making and using the invention is disclosed).

To the extent possible, please have them make any changes to the specification on this draft as they would like them to appear in the finalized application. Then when they have finished reviewing the application, please return the marked-up version of the draft. After we have received their comments and made any appropriate revisions, we will prepare the final application, which will be forwarded to you for your review.

If the application is in good order as enclosed or with the inventor's minor additions, please have the inventor's sign and date the enclosed declaration and assignment and return the originals to me for filing.

DOCKETED
By: [illegible]
Date: 4/24/01

As you are aware, we have a duty to disclose the most pertinent prior art of which you are aware to the Patent Office. If you can think of any pertinent references or patents, or simply any similar existing technology, please let us know, so that we can include a discussion of them with an Information Disclosure Statement. Additionally, please remember that the duty to disclose pertinent prior art continues until the patent actually issues. Therefore, if you become aware of other prior art designs in the future, please let us know.

Very truly yours,

Jessica

Jessica Pallach
Patent Administrator

/jlp

Enclosures